

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:52:43 ; Search time 53.2407 Seconds
(without alignments)
1302.928 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125

Perfect score: 125
Sequence: 1 TTGGCCCACTCCCTCTGCG.....CGCAGAGAGGAGTGGCCAA 125

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	145	US-07-789-917A-1	Sequence 1, Appli
2	125	100.0	145	US-08-702-573-4	Sequence 4, Appli
3	125	100.0	145	US-08-525-866-1	Sequence 1, Appli
4	125	100.0	145	US-07-982-193-1	Sequence 1, Appli
5	125	100.0	165	US-07-989-841A-1	Sequence 1, Appli
6	125	100.0	165	US-08-440-738A-1	Sequence 1, Appli
7	125	100.0	165	US-08-471-914-1	Sequence 1, Appli
8	125	100.0	165	US-09-276-625-7	Sequence 7, Appli
9	125	100.0	192	US-08-702-573-3	Sequence 3, Appli
10	125	100.0	4680	US-08-254-358-1	Sequence 1, Appli
11	125	100.0	4680	US-08-475-391-1	Sequence 1, Appli
12	125	100.0	4680	US-08-709-609-1	Sequence 1, Appli
13	125	100.0	4680	PCT-US95-07178-1	Sequence 1, Appli
14	125	100.0	5932	US-09-295-141-4	Sequence 4, Appli
15	125	100.0	5932	US-09-299-141-4	Sequence 4, Appli
16	125	100.0	6142	US-09-299-141-8	Sequence 8, Appli
17	125	100.0	6142	US-09-299-141-8	Sequence 8, Appli
18	125	100.0	6253	US-08-893-327-15	Sequence 15, Appli
19	125	100.0	6253	US-08-893-327-15	Sequence 15, Appli
20	125	100.0	6280	US-08-893-327-17	Sequence 17, Appli
21	125	100.0	6280	US-08-893-327-17	Sequence 17, Appli
22	125	100.0	6280	US-08-893-327-19	Sequence 19, Appli
23	125	100.0	6280	US-08-893-327-19	Sequence 19, Appli
24	125	100.0	6565	US-09-299-141-1	Sequence 1, Appli
25	125	100.0	6565	US-09-299-141-1	Sequence 1, Appli
26	125	100.0	6714	US-09-299-141-6	Sequence 6, Appli
27	125	100.0	6714	US-09-299-141-6	Sequence 6, Appli

28 125 100.0 6924 4 US-09-299-141-9 Sequence 9, Appli
29 125 100.0 6924 4 US-09-299-141-9 Sequence 9, Appli
30 125 100.0 6924 4 US-09-299-141-10 Sequence 10, Appli
31 125 100.0 6924 4 US-09-299-141-10 Sequence 10, Appli
32 125 100.0 6924 4 US-09-299-141-11 Sequence 11, Appli
33 125 100.0 6924 4 US-09-299-141-11 Sequence 11, Appli
34 125 100.0 6981 4 US-09-299-141-7 Sequence 7, Appli
35 125 100.0 6981 4 US-09-299-141-7 Sequence 7, Appli
36 125 100.0 7054 4 US-09-299-141-3 Sequence 3, Appli
37 125 100.0 7054 4 US-09-299-141-3 Sequence 3, Appli
38 125 100.0 7405 4 US-09-299-141-2 Sequence 2, Appli
39 125 100.0 7405 4 US-09-299-141-2 Sequence 2, Appli
40 125 100.0 7492 4 US-09-299-141-5 Sequence 5, Appli
41 125 100.0 7492 4 US-09-299-141-5 Sequence 5, Appli
42 125 100.0 8698 4 US-09-770-315-2 Sequence 2, Appli
43 123.4 98.7 272 4 US-09-276-625-4 Sequence 4, Appli
44 123 98.4 174 4 US-09-394-110A-1 Sequence 1, Appli
45 121.8 97.4 5585 2 US-08-305-221-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-07-789-917A-1
; Sequence 1, Application US/07789917A
; Patent No. 5252479
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Scully, Scott, Murphy Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release q.0, Version q.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07789,917A
; FILING DATE: 19911118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-789-917A-1

Query Match 100.0%; Score 125; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCCACTCCCTCTGCGCGCTCGCTGAGCGCGGCGGACCAAGGTGCGC 60
Db 1 TTGGCCCACTCCCTCTGCGCGCTCGCTGAGCGCGGCGGACCAAGGTGCGC 60
Qy 61 CGACGCCCGGGCTTTGCCCGGGCGGCGCTCAGTGAGCGCGCGGAGGAGTG 120

Db 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 2
US-08-702-573-4
; Sequence 4, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTA, Martine
; APPLICANT: DENEFE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: FERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..145
; OTHER INFORMATION: /note="Minimal ITR Sequence"
US-08-702-573-4

Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
QY 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGGAGTG 120
Db 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120

QY 121 GCCAA 125
Db 121 GCCAA 125
RESULT 3
US-08-525-866-1/c
; Sequence 1, Application US/08525866
; Patent No. 6207457
; GENERAL INFORMATION:
; APPLICANT: NATSOULIS, GEORGES
; APPLICANT: FUROSKY, RICHARD T.
; TITLE OF INVENTION: TARGETED NUCLEOTIDE SEQUENCE DELIVERY
; TITLE OF INVENTION: AND INTEGRATION SYSTEM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,866
; FILING DATE: 08-SEP-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 0800-0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-525-866-1

Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 125 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 66
QY 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGGAGTG 120
Db 65 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 6
QY 121 GCCAA 125
Db 5 GCCAA 1

RESULT 4
US-07-982-193-1
; Sequence 1, Application US/07982193
; Patent No. 6261834
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,193
; FILING DATE: 19921125
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-982-193-1
Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTGCGCGCTCGCTCACTGAGCGCGGCACCAAAGTCGCC 60
Db 1 TTGGCCACTCCCTCTGCGCGCTCGCTCACTGAGCGCGGCACCAAAGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGCCTCAGTGAGCGAGCGCGCAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGCCTCAGTGAGCGAGCGCGCAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 5
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match
Best Local Similarity 100.0%; Score 125; DB 2; Length 165;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 120
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 140

QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 7
US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samulski, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; FILE REFERENCE: 6636-027
; CURRENT APPLICATION NUMBER: US/08/471.914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440.738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: double-D
US-08-471-914-1

Query Match
Best Local Similarity 100.0%; Score 125; DB 3; Length 165;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 140

QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 8
US-09-276-625-7
; Sequence 7, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276.625
; CURRENT FILING DATE: 1999-03-25
```

```
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7

Query Match
Best Local Similarity 100.0%; Score 125; DB 4; Length 165;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGCGAGAGGGAGTG 140

QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 9
US-08-702-573-3
; Sequence 3, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTI, Martine
; APPLICANT: DENELE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: FERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702.573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..192
; OTHER INFORMATION: /note= "Right ITR Sequence in
; Patent No. 6033895
; OTHER INFORMATION: pXL2384"
US-08-702-573-3

Query Match      100.0%; Score 125; DB 3; Length 192;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTGCTCACTGAGCGCGGCGACCAAGTCCGC 60
Db 68 TTGGCCACTCCCTCTCTGCGGCTCGCTGCTCACTGAGCGCGGCGACCAAGTCCGC 127
Qy 61 CGACGCCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 128 CGACGCCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 187
Qy 121 GCCAA 125
Db 188 GCCAA 192

RESULT 10
US-08-254-358-1
; Sequence 1, Application US/08254358
; Patent No. 5658785
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; METHOD OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,358
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5658785and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Query Match      100.0%; Score 125; DB 1; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTGCTCACTGAGCGCGGCGACCAAGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTGCTCACTGAGCGCGGCGACCAAGTCCGC 60
Qy 61 CGACGCCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Qy 121 GCCAA 125
```

	DB	121 GCCAA 125		
RESULT 12				
US-08-709-609-1				
; Sequence 1, Application US/08709609				
; Patent No. 5858775				
; GENERAL INFORMATION:				
; APPLICANT: Johnson, Philip R.				
; TITLE OF INVENTION: Adeno-Associated Virus Materials and Methods				
; NUMBER OF SEQUENCES: 3				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun				
; STREET: 6300 Sears Tower, 233 S. Wacker Drive				
; CITY: Chicago				
; STATE: Illinois				
; COUNTRY: USA				
; ZIP: 60606				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: Patent In Release #1.0, Version #1.25				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/709,609				
; FILING DATE:				
; CLASSIFICATION: 435				
; ATTORNEY/AGENT INFORMATION:				
; NAME: No. 5858775and, Greta E.				
; REGISTRATION NUMBER: 35,302				
; REFERENCE/DOCKET NUMBER: 31975				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: (312) 474-6300				
; TELEFAX: (312) 474-0448				
; TELEX: 25-3856				
; INFORMATION FOR SEQ ID NO: 1:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 4680 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: DNA (genomic)				
; PCT-US95-07178-1				
; Query Match 100.0%; Score 125; DB 2; Length 4680;				
; Best Local Similarity 100.0%; Pred. No. 7.9e-25;				
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	TTGGCCACTCCCTCTGCGCGCTCGCTCACTAGAGCGCGGCCGACCAAAGTGC	60	
DB	1	TTGGCCACTCCCTCTGCGCGCTCGCTCACTAGAGCGCGGCCGACCAAAGTGC	60	
QY	61	CGAGCGCCGGGCTTTGCCGGCGGCTCAGTGACGCGCGCGAGAGGAGGTG	120	
DB	61	CGAGCGCCGGGCTTTGCCGGCGGCTCAGTGACGCGCGCGAGAGGAGGTG	120	
QY	121	GCCAA 125		
DB	121	GCCAA 125		
RESULT 14				
US-09-299-141-4				
; Sequence 4, Application US/09299141				
; Patent No. 6461606				
; GENERAL INFORMATION:				
; APPLICANT: FLOTTIE, TERENCE R.				
; APPLICANT: SONG, SIHONG				
; APPLICANT: BYRNE, BARRY J.				
; APPLICANT: MORGAN, MICHAEL				
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY				
; FILE REFERENCE: 4300.011800				
; CURRENT APPLICATION NUMBER: US/09/299,141				
; CURRENT FILING DATE: 1999-04-23				
; EARLIER APPLICATION NUMBER: 60/083,025				
; EARLIER FILING DATE: 1998-04-24				
; NUMBER OF SEQ ID NOS: 13				
; SOFTWARE: Patent In Ver. 2.0				
; SEQ ID NO 4				
; LENGTH: 5932				
; TYPE: DNA				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; OTHER INFORMATION: Description of Artificial Sequence:p43C-AT				
US-09-299-141-4				

Search completed: April 1, 2004, 17:04:17
Job time : 53.2407 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:05:53 ; Search time 1573.61 Seconds
(without alignments) 3442.959 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125
Perfect score: 125
Sequence: 1 TTGGCCACTCCCTCTCTGCG.....CGCAGAGGAGGAGGGCAA 125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	GenBank:
1: gb_ba.*	AX384518 Sequence
2: gb_hgt.*	AX384518 Sequence
3: gb_in.*	AX384518 Sequence
4: gb_ov.*	AX384518 Sequence
5: gb_ov.*	AX384518 Sequence
6: gb_ov.*	AX384518 Sequence
7: gb_ov.*	AX384518 Sequence
8: gb_ov.*	AX384518 Sequence
9: gb_ov.*	AX384518 Sequence
10: gb_ov.*	AX384518 Sequence
11: gb_ov.*	AX384518 Sequence
12: gb_ov.*	AX384518 Sequence
13: gb_ov.*	AX384518 Sequence
14: gb_ov.*	AX384518 Sequence
15: gb_ov.*	AX384518 Sequence
16: gb_ov.*	AX384518 Sequence
17: gb_ov.*	AX384518 Sequence
18: gb_ov.*	AX384518 Sequence
19: gb_ov.*	AX384518 Sequence
20: gb_ov.*	AX384518 Sequence
21: gb_ov.*	AX384518 Sequence
22: gb_ov.*	AX384518 Sequence
23: gb_ov.*	AX384518 Sequence
24: gb_ov.*	AX384518 Sequence
25: gb_ov.*	AX384518 Sequence
26: gb_ov.*	AX384518 Sequence
27: gb_ov.*	AX384518 Sequence
28: gb_ov.*	AX384518 Sequence
29: gb_ov.*	AX384518 Sequence
30: gb_ov.*	AX384518 Sequence
31: gb_ov.*	AX384518 Sequence
32: gb_ov.*	AX384518 Sequence
33: gb_ov.*	AX384518 Sequence
34: gb_ov.*	AX384518 Sequence
35: gb_ov.*	AX384518 Sequence
36: gb_ov.*	AX384518 Sequence
37: gb_ov.*	AX384518 Sequence
38: gb_ov.*	AX384518 Sequence
39: gb_ov.*	AX384518 Sequence
40: gb_ov.*	AX384518 Sequence
41: gb_ov.*	AX384518 Sequence
42: gb_ov.*	AX384518 Sequence
43: gb_ov.*	AX384518 Sequence
44: gb_ov.*	AX384518 Sequence
45: gb_ov.*	AX384518 Sequence

ALIGNMENTS

Result No.	Score	Query Match %	Length	DB ID	Description
1	125	100.0	130	6	AX384518 Sequence
2	125	100.0	145	6	AX384518 Sequence
3	125	100.0	145	6	AX384518 Sequence
4	125	100.0	145	6	AX384518 Sequence
5	125	100.0	145	14	AA2LTR1
6	125	100.0	145	14	AA2LTR2
7	125	100.0	145	14	AA2REPORI
8	125	100.0	165	6	AR034135
9	125	100.0	165	6	IL16806
10	125	100.0	165	6	AR223306
11	125	100.0	165	6	AX106702
12	125	100.0	165	6	BD218219
13	125	100.0	192	6	A46400
14	125	100.0	198	6	AX418199
15	125	100.0	207	6	AX703496
16	125	100.0	2116	14	AA2LEFT
17	125	100.0	4675	6	AX135805
18	125	100.0	4675	6	AX286292
19	125	100.0	4675	6	AX733252
20	125	100.0	4675	6	BD094552
21	125	100.0	4675	14	AA2CG
22	125	100.0	4679	6	AX282480
23	125	100.0	4679	14	AF043303
24	125	100.0	4680	6	AR028767
25	125	100.0	4680	6	162303
26	125	100.0	4681	6	BD242774
27	125	100.0	4683	6	BD242775
28	125	100.0	4683	14	AF028704
29	125	100.0	5932	6	AR235457
30	125	100.0	5932	6	AR235457
31	125	100.0	5932	6	AR235457
32	125	100.0	6142	6	AR235461
33	125	100.0	6565	6	AR235454
34	125	100.0	6565	6	AR235454
35	125	100.0	6714	6	AR235459
36	125	100.0	6714	6	AR235459
37	125	100.0	6924	6	AR235462
38	125	100.0	6924	6	AR235462
39	125	100.0	6924	6	AR235463
40	125	100.0	6924	6	AR235463
41	125	100.0	6924	6	AR235464
42	125	100.0	6924	6	AR235464
43	125	100.0	6981	6	AR235460
44	125	100.0	6981	6	AR235460
45	125	100.0	7054	6	AR235456

RESULT 1
AX384518
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX384518
Sequence 1 from Patent WO0214526.
AX384518
AX384518.1 GI:19577720
Adeno-associated virus 2
Adeno-associated virus 2
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
Li, S.
Replication competent aav helper functions
Patent: WO 0214526-A 1 21-FEB-2002;
Neurologix, Inc. (US)

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```

FEATURES
  source      Location/Qualifiers
    1..130
      /organism="Adeno-associated virus 2"
      /mol_type="unassigned DNA"
      /db_xref="taxon:10804"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 2
A46401
LOCUS      A46401      145 bp      DNA      linear      PAT 07-MAR-1997
DEFINITION      Sequence 4 from Patent WO9523867.
ACCESSION      A46401
VERSION      A46401.1 GI:2300602
KEYWORDS      .
SOURCE      unidentified
            unclassified.
REFERENCE      1 (bases 1 to 145)
AUTHORS      Densfle,P., Letta,M., Perricaudet,M. and Vigne,E.
TITLE      INTERACTIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND
            THERAPEUTICAL USES THEREOF
JOURNAL      Patente: WO 9523867-A 4 08-SEP-1995;
            RHONE POULENC RORER SA (FR)
COMMENT      Other publication AU 1852695 950918
            Other publication FR 2716893 950908.
FEATURES
  source      Location/Qualifiers
    1..145
      /organism="unidentified"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 3
AR140333/c
LOCUS      AR140333      145 bp      DNA      linear      PAT 16-JUN-2001
DEFINITION      Sequence 1 from patent US 6207457.
ACCESSION      AR140333
VERSION      AR140333.1 GI:14482829
KEYWORDS      .

```

```

SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 145)
AUTHORS      Natsoulis,G. and Surosky,R.T.
TITLE      Targeted nucleotide sequence delivery and integration system
JOURNAL      Patent: US 6207457-A 1 27-MAR-2001;
FEATURES
  source      Location/Qualifiers
    1..145
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 66
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
Db 65 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 6
QY 121 GCCAA 125
Db 5 GCCAA 1

RESULT 4
AX286293
LOCUS      AX286293      145 bp      DNA      linear      PAT 21-NOV-2001
DEFINITION      Sequence 2 from Patent WO0180840.
ACCESSION      AX286293
VERSION      AX286293.1 GI:17048541
KEYWORDS      .
SOURCE      Adeno-associated virus 2
            Adeno-associated virus 2
            Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE      1
AUTHORS      Raj,K. and Beard,P.M.
TITLE      Cytotoxic agents
JOURNAL      Patent: WO 0180840-A 2 01-NOV-2001;
            BTG INTERNATIONAL LIMITED (GB)
FEATURES
  source      Location/Qualifiers
    1..145
      /organism="Adeno-associated virus 2"
      /mol_type="unassigned DNA"
      /db_xref="taxon:10804"
      misc_structure 1..145 /notes="ITR"
      misc_feature 72
      misc_feature 94 /notes="Unpaired base"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
Db 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 140
QY 121 GCCAA 125
Db 141 GCCAA 145

```

```

RESULT 5
AA2LTR1      linear VRL 27-APR-1993
LOCUS        145 bp DNA
DEFINITION   Adeno-associated virus 2 left terminal sequence.
ACCESSION    K01624
VERSION      K01624.1 GI:209623
KEYWORDS     replication; terminal repeat.
SEGMENT      1 of 2
SOURCE       Adeno-associated virus 2H
ORGANISM     Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE    1 (bases 1 to 145)
AUTHORS      Lusby,E., Fife,K.H. and Berns,K.I.
TITLE        Nucleotide sequence of the inverted terminal repetition in
              adeno-associated virus DNA
JOURNAL      J Virol. 34 (2), 402-409 (1980)
MEDLINE      80185149
PUBMED       6246271
REFERENCE    2 (bases 1 to 145)
AUTHORS      Lefebvre,R.B., Riva,S. and Berns,K.I.
TITLE        Conformation takes precedence over sequence in adeno-associated
              virus DNA replication
JOURNAL      Mol. Cell. Biol. 4 (7), 1416-1419 (1984)
MEDLINE      85061247
PUBMED       6504049
COMMENT      Original source text: Adeno-associated virus 2H DNA, (clone pSM620
[2]), from KB or HeLa cells.
Both [1] and [2] present the opposite strand from the one presented
here. The focus of both papers is the method of replication of the
virus. [1] notes that the initial tt is present only 30% of the
time; it is shortened to t in 50% of the population and missing
altogether in 15% of the population. There is further sequence
heterogeneity which can be explained by assuming that the terminal
125 bases, which form an imperfect palindrome, are replaced by
their inverted complement during replication. [2] found that
deletion of the 9 terminal bases on the right and the 113 terminal
bases on the left of AAV 2 genome did not stop DNA replication.
Further deletion of an 11-base symmetrical sequence (bases 89 to
99) in the right terminal repetition inhibits DNA replication.
Substitution of either an 8-base (cagatctg) or 12-base
(cgcgatccgg) symmetrical sequence unrelated to the original
11-base sequence restores DNA replication. All of this can be
explained by assuming that the 125 base palindrome mentioned above
form a t-shaped secondary structure which provides a primer for DNA
polymerase during replication.
FEATURES     Location/Qualifiers
             1..145
                /organism="Adeno-associated virus 2H"
                /mol_type="genomic DNA"
                /db_xref="taxon:10805"
ORIGIN       2 bases upstream of HaeIII site.
Query Match          100.0%; Score 125; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY           1 TTGGCCACTCCCTCTGCGCGCTCGCTCACTGAGCGCGGCCGCACCAAGTGC GC 60
DB           1 TTGGCCACTCCCTCTGCGCGCTCGCTCGTCTACTGAGCGCGCGGCCGC 60
QY           61 CGAGCCCGGGCTTTGCCCGGGCGCGCTCACTGATGCGAGCGCGCAGAGGAGTG 120
DB           61 CGAGCCCGGGCTTTGCCCGGGCGCGCTCACTGATGCGAGCGCGCAGAGGAGTG 120
QY           121 GCCAA 125
DB           121 GCCAA 125

```

ORGANISM	Adeno-associated virus 2H
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
AUTHORS	1 (bases 1 to 145)
TITLE	Berns, K.I., Hauswirth, W.W., Fife, K.H. and Luseby, E.
JOURNAL	Adeno-associated virus DNA replication
MEDLINE	Cold Spring Harb. Symp. Quant. Biol. 43 Pt 2, 781-787 (1979)
PUBMED	80023388
COMMENT	226321
FEATURES	Original source text: Adeno associated virus 2H (AAV2 H) DNA. Location/Qualifiers 1..145 /organism="Adeno-associated virus 2H" /mol_type="genomic DNA" /db_xref="taxon:10805"

```

ORIGIN
Query Match      100.0%; Score 125; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY      1  TTGGCCACTCCCTCTCTCGGGCGCTCGCTCACTAGCGCCGGCGACCAAGGTCGCC 60
DB      145 TTGGCCACTCCCTCTCTCGGGCGCTCGCTCACTAGCGCCGGCGACCAAGGTCGCC 86
QY      61  CGACGCCCGGGGCTTTGCCCGGCGCGCTCACTAGCGCGAGCGAGCGCGCAGAGGGAGTG 120
DB      85  CGACGCCCGGGGCTTTGCCCGGCGCGCTCACTAGCGCGAGCGAGCGCGCAGAGGGAGTG 26
QY      121 GCCAA 125
DB      25 GCCAA 21

```

```

RESULT 8
AR034135
LOCUS
DEFINITION
AR034135
ACCESSION
AR034135.1
KEYWORDS
SOURCE
    Unknown.
ORGANISM
    Unclassified.
REFERENCE
    1 (bases 1 to 165)
AUTHORS
    Samulski, R. Jude. and Xiao, X.
TITLE
    Recombinant viral vector system
JOURNAL
    Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
    Location/Qualifiers
        1..165
            /organism="unknown"
            /mol_type="unassigned DNA"
source

```

```

/mot_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 60
      |||
Db      21  TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 80

QY      61  CGACGCCCGGGGTTTGCCCGCGGCGCTCACTGAGCGAGCGCGCGCAGAGGGAGTG 120
      |||
Db      81  CGACGCCCGGGGTTTGCCCGCGGCGCTCACTGAGCGAGCGCGCGCAGAGGGAGTG 140

QY      121 GCCAA 125
      |||
Db      141 GCCAA 145

```

RESULT	9				
116806					
LOCUS					
116806			165 bp	DNA	linear
					PAT 03-APR-1996

DEFINITION	Sequence 1 from patent US 5478745.
ACCESSION	I16806
VERSION	I16806.1
KEYWORDS	GI:1251714
SOURCE	.
ORGANISM	Unknown.
REFERENCE	Unknown.
AUTHORS	Unclassified.
JOURNAL	1 (bases 1 to 165)
FEATURES	Samulski, R.J. and Xiao, X. Title: Recombinant viral vector system Patent: US 5478745-A 1 26-DEC-1995; Location/Qualifiers 1..165 /organism="unknown" /mol type="unassigned DNA"
source	

```

ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2,7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTGGCCACTCCCTCTCTGGCGGCTCGCTGCTCACTGAGGCCGGGGCGACCAAGGTCGCC 60
Db      21  TTGGCCACTCCCTCTCTGGCGGCTCGCTGCTCACTGAGGCCGGGGCGACCAAGGTCGCC 80

QY      61  CGAGCGCCGGGGCTTTGCGCCGGGGCGGCTCAGTCAGCGAGCGAGCGCGCAGAGAGGGAGTG 120
Db      81  CGAGCGCCGGGGCTTTGCGCCGGGGCGGCTCAGTCAGCGAGCGAGCGCGCAGAGAGGGAGTG 140

QY      121 GCCAA 125
Db      141 GCCAA 145

```

```

RESULT 10
AR223306 LOCUS linear
DEFINITION Sequence 7 from patent US 6436392.
ACCESSION AR223306
VERSION AR223306.1 GI:23331457
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 165)
AUTHORS Engelhardt,J.F. and Duan,D.
TITLE Adeno-associated virus vectors
JOURNAL Patent: US 6436392-A 7 20-AUG-2002;
FEATURES
    Location/Qualifiers
        1..165
            /organism="unknown"
            /mol_type="genomic DNA"
    source

```

```

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2,7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/ mol_type="genomic DNA"

Qy 1 TTGGCACTCCCTCTCTGGCGCTCGCTCTACTGAGCGCGGGACCAAAGGTGCC 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21 TTGGCACTCCCTCTCTGGCGCTCGCTCTACTGAGCGCGGGACCAAAGGTGCC 80
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 140
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 GCCAA 125
| | | | |
Db 141 GCCAA 145
| | | | |

```

RESULT 11
AX106702

```
LOCUS AX106702 165 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0125465.
ACCESSION AX106702
VERSION AX106702.1 GI:13922363
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Engelhardt, J.F., Dongsheng, D. and Ziyang, Y.
TITLE Adeno-associated viruses and uses thereof
JOURNAL Patent: WO 0125465-A 7 12-APR-2001;
University of Iowa Research Foundation (US); Engelhardt, John F.
(US); Dongsheng, Duan (US); Ziyang, Yan (US)
FEATURES
source
1..165
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="SEQ ID NO:1 of U.S. Patent No. 5,478,745"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTAGCGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTAGCGCGCGGCGACCAAGGTCGCC 80

QY 61 CGAGCCCGGGCTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCGAGAGGAGTG 120
DB 81 CGAGCCCGGGCTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 13
A46400
LOCUS A46400 192 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 3 from Patent WO9523867.
ACCESSION A46400
VERSION A46400.1 GI:2300601
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 192)
AUTHORS Densfle, P., Latta, M., Perricaudet, M. and Vigne, B.
TITLE INTERACTIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
JOURNAL Patent: WO 9523867-A 3 08-SEP-1995;
PHONE POULENC RORER SA (FR)
COMMENT Other publication AU 1852695 950918
Other publication FR 2716893 950908.
FEATURES
source
1..192
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTAGCGCGCGGCGACCAAGGTCGCC 60
DB 68 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTAGCGCGCGGCGACCAAGGTCGCC 127

QY 61 CGAGCCCGGGCTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCGAGAGGAGTG 120
DB 128 CGAGCCCGGGCTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCGAGAGGAGTG 187

QY 121 GCCAA 125
DB 188 GCCAA 192

RESULT 14
AX418199/c
LOCUS AX418199/c 198 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0192551.
ACCESSION AX418199
VERSION AX418199.1 GI:21523203
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Samulski, R.J. and Mccarty, D.M.
TITLE Duplexed parvovirus vectors
```

JOURNAL Patent: WO 0192551-A 1 06-DEC-2001;
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US)

FEATURES
source
1. .198
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Inverted terminal repeat from the AAV-2 vector
plasmid pSub 201"

ORIGIN

Query Match 100.0%; Score 125; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGC 60
Db 125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGC 66
QY 61 CGACGCCCGGGGTTTGGCGCGCGGCTCAGTGAGCGAGCGCGCGGAGGAGTG 120
Db 65 CGACGCCCGGGGTTTGGCGCGCGGCTCAGTGAGCGAGCGCGCGGAGGAGTG 6
QY 121 GCCAA 125
Db 5 GCCAA 1

RESULT 15

AX703496
LOCUS AX703496 207 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 58 from Patent WO02066653.
ACCESSION AX703496
VERSION AX703496.1 GI:29538456
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Li, M. and Liu, Y. C.
TITLE Prokaryotic libraries and uses
JOURNAL Patent: WO 02066653-A 58 29-AUG-2002;
Xencor (US)

FEATURES

source
1. .207
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic enzyme attachment site sequence"

ORIGIN

Query Match 100.0%; Score 125; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGC 60
Db 42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGC 101
QY 61 CGACGCCCGGGGTTTGGCGCGCGGCTCAGTGAGCGAGCGCGCGGAGGAGTG 120
Db 102 CGACGCCCGGGGTTTGGCGCGCGGCTCAGTGAGCGAGCGCGCGGAGGAGTG 161
QY 121 GCCAA 125
Db 162 GCCAA 166

Search completed: April 1, 2004, 16:04:10
Job time : 1574.61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:50:59 ; Search time 1603.24 Seconds
(without alignments)
2328.267 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125
Perfect score: 125
Sequence: 1 TTGGCCACTCCTCTCTGCG.....CGCAGAGAGGAGTGGCCAA 125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_esth.*

9: gb_est1.*

10: gb_est2.*

11: gb_estc.*

12: gb_estd.*

13: gb_este.*

14: gb_estf.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_nam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_pbg.*

27: em_gss_vri.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	30.2	304	10	B2495968
C 2	37.4	29.9	1015	28	B2569259
C 3	37	29.6	1049	29	AG185368
C 4	36.6	29.3	644	13	BY752719

5	35.4	28.3	1103	13	BX403654
6	35.4	28.3	1182	12	BM807831
C 7	35	28.0	595	28	BH218838
C 8	35	28.0	896	13	BU152072
C 9	35	28.0	1103	13	BX403654
C 10	34.8	27.8	530	12	BMQ79047
C 11	34.6	27.7	514	13	BX338963
C 12	34.6	27.7	598	29	CG273564
C 13	34.6	27.7	654	28	B2532489
C 14	34.6	27.7	747	28	B2410270
C 15	34.6	27.7	827	29	CG285957
C 16	34.6	27.7	839	29	CG365468
C 17	34.6	27.7	840	29	CG441848
C 18	34.6	27.7	932	29	CNS00720
C 19	34.4	27.5	878	9	AL552821
C 20	34.2	27.4	969	9	AL564360
C 21	34	27.2	640	13	BY733594
C 22	34	27.2	933	29	AG060831
C 23	34	27.2	1121	29	CNS05GBJ
C 24	34	27.2	1123	29	AG080476
C 25	34	27.2	1201	13	BX405071
C 26	33.8	27.0	663	14	CB650788
C 27	33.8	27.0	1000	13	BX407619
C 28	33.8	27.0	1697	10	AW727558
C 29	33.6	26.9	430	29	CT78537
C 30	33.6	26.9	1057	29	CNS046BO
C 31	33.6	26.9	1057	29	CNS046BO
C 32	33.6	26.9	1471	28	BZ571163
C 33	33.4	26.7	433	14	CA697866
C 34	33.4	26.7	509	29	CG712785
C 35	33.4	26.7	514	13	BX338963
C 36	33.4	26.7	574	28	BH225085
C 37	33.4	26.7	578	28	BH218943
C 38	33.4	26.7	579	28	BH218776
C 39	33.4	26.7	594	28	BH218854
C 40	33.4	26.7	630	29	CNS01PKV
C 41	33.4	26.7	808	28	BZ742499
C 42	33.4	26.7	1011	13	BQ961334
C 43	33.4	26.7	1049	12	BQ048852
C 44	33.4	26.7	1201	13	BX403541
C 45	33.2	26.6	366	13	BY303434

ALIGNMENTS

RESULT 1	BE495968/c	304 bp	mRNA	linear	EST 02-AUG-2000
LOCUS	WHE1260_H03_006ZS	Secale cereale	anther cDNA library	Secale cereale	
DEFINITION	CDNA clone WHE1260_H03_006, mRNA sequence.				
ACCESSION	BE495968				
VERSION	EST.				
KEYWORDS	EST.				
SOURCE	Secale cereale (rye)				
ORGANISM	Secale cereale				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Poaceae; Secale.				
AUTHORS	Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.I. and Tong, J.C.				
TITLE	The structure and function of the expressed portion of the wheat				
JOURNAL	Genomes - Another cDNA library from rye				
COMMENT	Unpublished (2000)				
	Contact: Olin Anderson				
	US Department of Agriculture, Agriculture Research Service, Pacific				
	West Area, Western Regional Research Center				
	800 Buchanan Street, Albany, CA 94710, USA				
	Tel: 5105595773				
	Fax: 5105595818				


```

QY 64 CGCCCGGCTTTCCCGGGGGCTCAGTGGCGAGCGAGCGCGGAGAGGAGTGG 121
    |||||
Db 467 CGCCCGGCTTTCCCGGGGGCTCAGTGGCGAGCGAGCGCGGAGAGGAGTGG 410
    |||||

RESULT 4
BY752719
LOCUS BY752719 RIKEN full-length enriched, adult inner ear Mus musculus
DEFINITION cDNA clone F930013F16 5', mRNA sequence.
ACCESSION BY752719
VERSION BY752719.1 GI:27183756
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 644)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Schonbach,C., Gojobori,T., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.B., Cousins,S., Dalia,E., Dragan,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierzki,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wyshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
22354683
12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/

ADACHI,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,M., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format

```

```

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Kirk W. Beisel ( Boys Town National
Research Hospital 555 North 30th Street Omaha,NE 68131 USA ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```

FEATURES

source

```

1..644
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930013F16"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"

```

ORIGIN

```

Query Match      29.3%; Score 36.6; DB 13; Length 644;
Best Local Similarity 56.1%; Pred. No. 84;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 3  GGCCACTCCCTCTCTGCGGCTCGCTCAGTGGCGAGCGCGGAGGAGTGGCGG 62
    |||||
Db 217 GGCCACTCCGCTATTTGGGCTTGTCTGTGCGCTGGAGCGCGGCGGCGGCGT 276

QY 63  AGCCCGGCGCTTTGCCCGGGCGCTCAGTGGCGAGCGCGGAGGAGTGGCG 122
    |||||
Db 277 CCCCACACCTGTCTCCAGGCTGCGCTGGAGCGCGGCGGCGGCGGCGTGC 336

QY 123 CAA 125
    ||
Db 337 AAA 339

```

RESULT 5

BX403654

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

```

1..1103
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

BX403654 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA007ZH12
3-PRIME, mRNA sequence.
BX403654.1 GI:30762430
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1103)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BA007ZH12FP1.

```


Db	538	GGCGCCCTCCCTCTCCGCGGAGGCGGCTCCCTGAGGACGTGGTGC	475
Qy	64	CGCCCGGGCTTTGCCCGGGCGGGCCCTAGTGAGCGAGCGCGGCCAGAGAGGAG	118
Db	478	CGGTGGCTTCTGTCCCGAAGCGGGGGTTTCGGGGCGACCTCGAGATGAACGAG	424


```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI065YP08"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (GT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

ORIGIN
Query Match          27.7%  Score 34.6;  DB 13;  Length 514;
Best Local Similarity 28.9%  Pred. No. 2.4e-02;
Matches 35;  Conservative 43;  Mismatches 43;  Indels 0;  Gaps 0;

sites of the pCMVSPORT 6 vector. Library was normalized.

QY 5  CCACTCCCTCTTCGCGCTGCTCCTCACTGAGCGCGGCGACCAAGATCGCCGAC 64
Db 471  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 412
QY 65  GCCCGGGGTTTCCCGGGCGGCTCTCAGTGAGCGAGCGCGAGAGAGGAGTGCGCA 124
Db 411  SCSSSSSCSSSSSCSSCCSCCCSCSSSSSSSCSCSSSSSSSSSSSSSSSSSSSS 352
QY 125  A 125
Db 351  R 351

```

DEFINITION	CG2BBL16TH.ZM.0.7.1.5.KB Zea mays genomic clone ZMMBMA0750D08, genomic survey sequence.
ACCESSION	CG273564
VERSION	CG273564.1
KEYWORDS	GI:34185705
SOURCE	GSS.
ORGANISM	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 598)
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
TITLE	Consortium for Maize Genomics
JOURNAL	Unpublished (2002)
COMMENT	Contact: Cathy Whitelaw

```

JOURNAL COMMENT
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0286
Email: whitelaw@cigr.org
Seq primer: TK
Class: sheared ends..
Location/Qualifiers
1..598
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0750D08"
/clone.lib="ZM 0.7 1.5 KB"
note="vector: pCSCS-; Site 1: HincII; 0.7-1.5 kb
maturation filtered genomic DNA library"
FEATURES
source

```

QY 18 GCGGGCTCGCTCGCTCACTGAGCCCGCGCACCAAGGTGCGCCCGAGCGCCCGGGCTTTGC 77

```

Db 169 GTGAGCTTGGTGAAGCACTTGAGCGGCGCAACCTCCGCCAGTATGGAGCAGGAGTGGC 228
QY 78 CGCGGCGGCTCAGTGAGCGAGCGAGCGCGCGAGAG 114
Db 229 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 265

RESULT 13
BZ532489
LOCUS OGAEY44TC ZM2.0.7.1.5_KB Zea mays genomic clone ZM5BMA0046G15,
DEFINITION genomic survey sequence.
ACCESSION BZ532489
VERSION BZ532489.1 GI:27077191
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 654)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..654
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0046G15"
/clone_lib="ZM2_0.7.1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
1..654
Query Match 27.7%; Score 34.6; DB 28; Length 654;
Best Local Similarity 59.8%; Pred No. 2.5e-02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGCGGCGCAAAAGTGCAGCCCGGGCTTTGC 77
Db 513 GTGAGCTTGGTGAAGCACTTGAGCGGCGCAACCTCCGCCAGTATGGAGCAGGAGTGGC 572
QY 78 CGCGGCGGCTCAGTGAGCGAGCGCGCGAGAG 114
Db 573 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 609

RESULT 14
BZ410270/c
LOCUS OGAB801TC ZM0.7.1.5_KB Zea mays genomic clone ZM5BMA0020023,
DEFINITION genomic survey sequence.
ACCESSION BZ410270
VERSION BZ410270.1 GI:26042544
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 747)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..827
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA00607P17"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
1..827
Query Match 27.7%; Score 34.6; DB 28; Length 747;
Best Local Similarity 59.8%; Pred No. 2.6e-02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGCGGCGCAAAAGTGCAGCCCGGGCTTTGC 77
Db 593 GTGAGCTTGGTGAAGCACTTGAGCGGCGCAACCTCCGCCAGTATGGAGCAGGAGTGGC 534
QY 78 CGCGGCGGCTCAGTGAGCGAGCGCGCGAGAG 114
Db 533 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 497

RESULT 15
CG285957/c
LOCUS OGWK393TH ZM0.7.1.5_KB Zea mays genomic clone ZM5BMA0607P17,
DEFINITION genomic survey sequence.
ACCESSION CG285957
VERSION CG285957.1 GI:34200171
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 827)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..827
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0607P17"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
1..827
Query Match 27.7%; Score 34.6; DB 28; Length 747;
Best Local Similarity 59.8%; Pred No. 2.6e-02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGCGGCGCAAAAGTGCAGCCCGGGCTTTGC 77
Db 593 GTGAGCTTGGTGAAGCACTTGAGCGGCGCAACCTCCGCCAGTATGGAGCAGGAGTGGC 534
QY 78 CGCGGCGGCTCAGTGAGCGAGCGCGCGAGAG 114
Db 533 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 497

RESULT 15
CG285957/c
LOCUS OGWK393TH ZM0.7.1.5_KB Zea mays genomic clone ZM5BMA0607P17,
DEFINITION genomic survey sequence.
ACCESSION CG285957
VERSION CG285957.1 GI:34200171
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 827)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..827
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM5BMA0607P17"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
1..827
Query Match 27.7%; Score 34.6; DB 28; Length 747;
Best Local Similarity 59.8%; Pred No. 2.6e-02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGCGGCGCAAAAGTGCAGCCCGGGCTTTGC 77
Db 593 GTGAGCTTGGTGAAGCACTTGAGCGGCGCAACCTCCGCCAGTATGGAGCAGGAGTGGC 534
QY 78 CGCGGCGGCTCAGTGAGCGAGCGCGCGAGAG 114
Db 533 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 497

```

/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match	27.7%	Score 34.6	DB 29	Length 827
Best Local Similarity	59.8%	Pred. No. 2.7e+02		
Matches	58	Conservative	0	Mismatches 39
				Indels 0
				Gaps 0

QY	18	GCGCGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCCCGCCGCGGGGCTTGC	77
Db	351	GTGAGCTTGGTGAGGCACTTGAGCCGGGCAACCTCCGCCAGTATGGAGCAGGATGCG	292

QY	78	CCGGGGGGCTCAGTGAGCGAGCGCGCGCAGAGAG	114
Db	291	CTGGGGGGCACCATTGAGAGCTCGAGCTCCTTGAGAG	255

Search completed: April 1, 2004, 17:02:17
Job time : 1606.24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:50:59 ; Search time 1859.76 Seconds
(without alignments)
2328.267 Million cell updates/sec

Title: US-10-620-039-1
Perfect score: 145
Sequence: 1 TTGCCACTCTCTCTGCG.....CTCCATCACTAGGGGTTTCCT 145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estm:*
 - 5: em_estcov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rtd:*
 - 26: em_gss_pig:*
 - 27: em_gss_vrl:*
 - 28: gb_ges1:*
 - 29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	38.6	26.6	974	14	CF784470
C 2	37.8	26.1	304	10	BE495968
C 3	37.4	25.8	1015	28	BZ569259
C 4	37	25.5	1049	29	AG185368

5	36.6	25.2	644	13	BY752719
6	36.4	25.1	969	9	AL564360
C 7	35.8	24.7	1121	29	CNS05GBJ
C 8	35.6	24.6	1103	13	AX403654
9	35.6	24.6	1201	13	AX405071
10	35.4	24.4	598	28	CG273564
11	35.4	24.4	554	28	BZ53489
C 12	35.4	24.4	747	28	BZ410270
C 13	35.4	24.4	827	29	CG285957
14	35.4	24.4	839	29	CG365468
15	35.4	24.4	840	29	CG441848
16	35.4	24.4	1103	13	AX403654
17	35.4	24.4	1182	12	BM807831
18	35.2	24.3	1349	11	AX017628
C 19	35	24.1	595	28	BZ18838
20	35	24.1	896	13	BUI52072
21	35	24.1	1303	28	BZ570288
C 22	34.8	24.0	530	12	BM079047
C 23	34.8	24.0	953	13	BQ225579
C 24	34.6	23.9	514	13	BX338963
C 25	34.6	23.9	813	28	AZ193936
C 26	34.6	23.9	932	29	CNS0072Q
C 27	34.4	23.7	878	9	AL552621
C 28	34.2	23.6	1049	12	BQ048852
C 29	34.2	23.6	1201	13	BX360624
C 30	34	23.4	438	12	BM500627
31	34	23.4	640	13	BY733594
32	34	23.4	933	29	AG060831
33	34	23.4	1000	13	BX407619
34	34	23.4	1101	29	CNS017V2
35	34	23.4	1115	12	BM924606
36	34	23.4	1123	29	AG080476
37	33.8	23.3	663	14	CD650788
C 38	33.8	23.3	927	28	CG390088
39	33.8	23.3	1057	29	CNS046BO
40	33.8	23.3	1697	10	AW727558
41	33.6	23.2	430	29	CC778537
C 42	33.6	23.2	650	12	B1549023
C 43	33.6	23.2	874	13	BX389389
C 44	33.6	23.2	891	12	B1551983
C 45	33.6	23.2	1057	29	CNS046BO

ALIGNMENTS

RESULT 1	CF784470/c	CF784470	774 bp	mrna	linear	EST 20-OCT-2003
LOCUS	AGENCOURT_15918074	X-St10-30	Silurana tropicalis	cDNA clone		
DEFINITION	IMAGE:7027270 3', mRNA sequence.					
ACCESSION	CF784470					
VERSION	CF784470.1	GI:37748315				
KEYWORDS	EST.					
SOURCE	Silurana tropicalis (western clawed frog)					
ORGANISM	Silurana tropicalis					
REFERENCE	1 (bases 1 to 974)					
AUTHORS	NIH-MGC	http://imgc.ncbi.nih.gov/				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgabs-r@mail.nih.gov Tissue Procurement: Dr. Wei Wu / Prof. Christof Niehrs CDNA Library Preparation: Wei Wu, DKPZ CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be					

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14761 row: d column: 20

High quality sequence start: 36

High quality sequence stop: 157

Location/Qualifiers

1..974

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone_image="7027270"

/tissue_type="whole embryo, pool of stages 10, 20 and 30"

/note="Vector: PRW2; Site 1: BamHI; Site 2: XhoI; 10 ug

of polyA+ RNA was isolated from a mixture of embryos at

stage 10, 20 and 30 and primed by oligo-dT primer:

5'-GAGAGAGAGAGATCC(T)16VN-3' (where V=G,A,C)

5-methyl-dCTP was used instead of dCTP in the first-strand

synthesis in order to get hemimethylated cDNA. After

full-length enrichment, oligo-dG tailing and normalization

against itself, second-strand synthesis was carried out by

priming with 5'-GAGAGAGAGAGTAAATTAAT(C)13-3'. dsDNA

was digested with XhoI/BamHI and directionally cloned into

the pRW2 vector. Average insert size is 1.5 kb. Library

constructed using the Carninci protocol (Genome Research

2000) by Drs. W. Wu and C. Niehrs (DFPZ, Heidelberg,

Germany)."

ORIGIN

Query Match 26.6%; Score 38.6; DB 14; Length 974;
Best Local Similarity 91.1%; Pred. No. 35;
Matches 41; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 101 GAGCGCGAGAGGGGCGCACTCCATCCTAGGGGTTCT 145

Db 159 GTGATTGAGAGGGGAGGCGCACTCCATCCTAGGGGTTCT 115

RESULT 2

BE495968/c

LOCUS

DEFINITION WHE1260_H03 006ZS Secale cereale anther cDNA library Secale cereale

cDNA clone WHE1260_H03_006, mRNA sequence.

BE495968

VERSION

KEYWORDS

SOURCE

ORGANISM

Secale cereale (rye)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Secale.

1 (bases 1 to 304)

Anderson, O.D., Butler, E., Chao, S., Choi, D.W., Close, T.J.,

Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y.,

Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and

Tong, J.C.

The structure and function of the expressed portion of the wheat

Genomes - Anther cDNA library from rye

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105395773

Fax: 5105395818

Email: andersowp.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

FEATURES

source

1..304

/organism="Secale cereale"

/mol_type="mRNA"

/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1260_H03_006"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. coli SOLR"
/clone_lib="Secale cereale anther cDNA library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse. Anthers were harvested and pooled from early
meiosis to late meiosis. The tissue, total RNA, and
poly(A) RNA were prepared (Butler, Ross and Gustafson) at
University of Missouri, Columbia. A cDNA library was
made, and the cDNA clones were in vivo excised to give
pBluescript phagemids in the T3 Close lab (Choi, Close,
Fenton) at the University of California, Riverside.
Plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 26.1%; Score 37.8; DB 10; Length 304;
Best Local Similarity 57.9%; Pred. No. 38;
Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 2 TGGCCACTCCCTCTCTGCGGCTCGCTCAGTGGCGGCGACCAAGGTGCCCC 61

Db 118 TTGCCATCTCTTCTCAGAGCTCACTCAGGACAGAGGTAGGAGAGACACACTA 59

QY 62 GAGCCCGGCTTTGCCGGCGGCTCAGTGGCGGCGGCGGCGGCGGCGGAGG 115

Db 58 GACACAGAGGATTTTCCGGNCCGCGCACAGACCGCGGTGGGCGCAGACG 5

RESULT 3

BZ569259/c

LOCUS

DEFINITION BZ569259 1015 bp DNA linear GSS 17-DEC-2002

clone pacs2-164_8319.x2r5p12 pacs2-164 Pseudomonas aeruginosa genomic

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 1015)

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

Burns, J.L., Kaul, R. and Olsen, M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: Shotgun.

Location/Qualifiers

1..1015

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="2-164"

/db_xref="taxon:287"

/clone_lib="pacs2-164_8319"

/notes="clinical isolate 2-164 Whole genomic shotgun

library."

ORIGIN

Query Match 25.8%; Score 37.4; DB 28; Length 1015;

Best Local Similarity 56.7%; Pred. No. 68;

```

Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
Qy 4 GCACCTCCTCTCTGCGCGTCTGCTCCTCAGTGGCCGGCGACCAAGGTCGCCGGA 63
Db 952 GCCTCGCGCGTGGCGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 893
Qy 64 CGCCCGGCGCTTTGCCCGCGCGCTCAGTGGAGCGAGCGCGCGAGAGGAGGAGTGCC 123
Db 892 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 833

RESULT 4
AG185368/c
LOCUS
DEFINITION
  Pan troglodytes DNA, clone: RP43-059G23.TJ, genomic survey
  sequence.
ACCESSION
  AG185368
VERSION
  AG185368.1 GI:16715048
KEYWORDS
  GSS.
ORGANISM
  Pan troglodytes (chimpanzee)
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Mammalia; Euthera; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
  1 Fujiyama A., Hattori M., Toyoda A., Taylor T.D., Yada T.,
  Totoki Y., Watanabe H. and Sakaki Y.
  BAC end sequences of Library RPCI-43
  Unpublished
REFERENCE
  2 (bases 1 to 1049)
  Fujiyama A., Hattori M., Toyoda A., Taylor T.D., Yada T.,
  Totoki Y., Watanabe H. and Sakaki Y.
  Direct Submission
TITLE
  Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
  and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  (E-mail: chimpanzee@gscc.riken.go.jp, URL: http://hgpc.gsc.riken.go.jp/,
  Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT
  Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
  end was generated during the R&D process and may have higher chance
  of clone tracking errors.
PRIMERS
  Sequencing: TJ
LIBRARY
  Vector : pBACe3.6
  R.Site 1 : EcoRI
  R.Site 2 : EcoRI
  Location/Qualifiers
    1..1049
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="RP43-059G23.TJ"
    /sex="male"
    /cell_type="lymphocytes"
    /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN
  Query Match 25.5%; Score 37; DB 29; Length 1049;
  Best Local Similarity 56.8%; Pred. No. 86;
  Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
  Qy 4 GCACCTCCTCTCTGCGCGTCTGCTCCTCAGTGGCCGGCGACCAAGGTCGCCGGA 63
  Db 527 GCCTCGCGCGTGGCGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 468
  Qy 64 CGCCCGGCGCTTTGCCCGCGCGCTCAGTGGAGCGAGCGCGCGAGAGGAGGAGTG 121
  Db 467 CGCCCGCGCGCGCGCGCGCGTCTCGCGCTGCGCGCGCGCGCGCGCGCGTGG 410

RESULT 5
BY752719
LOCUS

```

```

DEFINITION
  BY752719 RIKEN full-length enriched, adult inner ear Mus musculus
  cDNA clone F330013F16 5', mRNA sequence.
ACCESSION
  BY752719
VERSION
  BY752719.1 GI:27183756
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 644)
  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
  Nikaido I., Oeato N., Saito R., Suzuki H., Yamataka I.,
  Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A.,
  Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C.,
  Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H.,
  Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V.,
  Chothia C., Corbani L.E., Cousins S., Della E., Dragani T.A.,
  Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T.,
  Gariboldi M., Glessi C., Godzik A., Gough J., Grimmond S.,
  Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A.,
  Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A.,
  Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R.,
  Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T.,
  Nunata K., Okido T., Pavan W.J., Pertea G., Peocle G.,
  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
  Ravasi F., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
  Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
  Verdaro R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y.,
  Wells C., Wilming L.G., Wyntshaw-Boris A., Yanagisawa M., Yang I.,
  Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P.,
  Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M.,
  Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K.,
  Aizawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y.,
  Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K.,
  Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S.,
  Rogers J., Birney E. and Hayashizaki Y.
TITLE
  Analysis of the mouse transcriptome based on functional annotation
  of 50,770 full-length cDNAs
JOURNAL
  Nature 420, 563-573 (2002)
MEDLINE
  22354683
PUBMED
  12466851
COMMENT
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center (GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gscc.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/
  Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P.,
  Fukuda S., Hashizume W., Hayashida K., Hirozane T., Hori F.,
  Imotani K., Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y.,
  Kondo S., Konno H., Koya S., Miyazaki A., Murata M., Nakamura M.,
  Nomura K., Numazaki R., Ohno M., Ohsato N., Saito R., Sakazume N.,
  Sano H., Sasaki D., Sato K., Shibata K., Shiraki T., Tagami M.,
  Takeda Y., Waki K., Watanabe A., Muramatsu M. and Hayashizaki Y.
  Direct Submission
  Computational Analysis of Full-length Mouse cDNAs Compared with
  Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Computer-based methods for the mouse full-length cDNA
  encyclopedia: real-time sequence clustering for construction of a
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.

```



```
/clone="003C03"
/clone.lib="B"
/note="Genoscope sequence ID : COAB003AB02C1-end : T7"

ORIGIN

Query Match      24.7%; Score 35.8; DB 29; Length 1121;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 64; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 GCCACTCCTCTCTGCGGCTCGTCTACTGAGCGCGGCGACCAAGGTGCGCCGA 63
DB 664 GCGCGCNCNNGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 605

QY 64 GCGCGCGGCTTTGCGCGGCGCGCTCTAGTAGCGAGCGCGCGCGCGAGAGGAGTGCC 123
DB 604 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545

QY 124 AACTCCATCACTAGG 139
DB 544 GCGCNCNCGCGCGCG 529

RESULT 8
BX403654/c
LOCUS BX403654 Homo sapiens PLACENTA mRNA linear EST 15-MAY-2003
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX403654
VERSION BX403654.1 GI:30762430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (Bases 1 to 1103)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA007ZH12FP1.
Location/Qualifiers
1..1103
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA007ZH12"
/tissue_type="PLACENTA"
/clone.lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match      24.6%; Score 35.6; DB 13; Length 1103;
Best Local Similarity 12.3%; Pred. No. 1.9e+02;
Matches 15; Conservative 71; Mismatches 36; Indels 0; Gaps 0;

QY 13 TCTCTGCGGCTCGCTCCTACTGAGCGCGGCGACCAAGGTGCGCGCGCGCGCG 72
DB 503 WSTCTSSSSSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 444

QY 73 TTGCGCGGCGGCTCTAGTAGCGAGCGCGCGCGAGAGGAGTGCCCACTCCATC 132
DB 443 SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 384
```

```
QY 133 AC 134
DB 383 CC 382

RESULT 9
BX405071
LOCUS BX405071 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0D006YG06 3-PRIME, mRNA sequence.
ACCESSION BX405071
VERSION BX405071.1 GI:30648111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10245.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AG006BD03NP1&cluster=10245.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AG006BD03NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D006YG06"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/clone.lib="RAMOS CELL LINE"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match      24.6%; Score 35.6; DB 13; Length 1201;
Best Local Similarity 43.8%; Pred. No. 1.9e+02;
Matches 56; Conservative 19; Mismatches 53; Indels 0; Gaps 0;

QY 3 GGCCTCTCTCTGCGGCTCGCTCTACTGAGCGCGGCGACCAAGGTGCGCCG 62
DB 741 GGCACMGACACCTCTCGCGCGCGCGCGSSCGCGCGCGCGCGCGCGCGCGCG 800

QY 63 AGCCCGCGGCTTTGCGCGGCGCGCTCTAGTAGCGAGCGCGCGAGAGAGTGAGTGGC 122
DB 801 SCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860

QY 123 CAATCTCA 130
DB 861 TGGCGCCA 868

RESULT 10
CG273564
LOCUS CG273564 598 bp DNA linear GSS 25-AUG-2003
DEFINITION CG2B16TH_ZM_0.7_1.5_KB zea may genomic clone ZMBEma0750D08,
genomic survey sequence.
ACCESSION CG273564
VERSION CG273564.1 GI:34185705
```

```

KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 598)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source
1..598
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0750D08"
/notes="Vector: pBCSK-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 24.4%; Score 35.4; DB 29; Length 598;
Best Local Similarity 57.8%; Pred.No.1.7e+02; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 46;
QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCCGACGCCCGGGCTTTGC 77
Db 169 GTGAGCTTGGTGAGGCACCTTGAGCGGGCAACCTCGCCAGTATGAGCAGGAGTGCC 228
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTGCCCAAC 126
Db 229 CTGGGGCGGCACCATGAGAGCTCGAGCTCCTTGAGAGCGCTGCATCCGAC 277

RESULT 11
BZ532489
LOCUS
DEFINITION
CGAEY447C ZM2.0.7.1.5_KB Zea mays genomic clone ZMMEMa0046G15,
genomic survey sequence.
ACCESSION
BZ532489
VERSION
BZ532489.1 GI:27077191
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 654)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source
1..654
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0046G15"
/notes="Vector: pBCSK-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 24.4%; Score 35.4; DB 29; Length 598;
Best Local Similarity 57.8%; Pred.No.1.7e+02; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 46;
QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCCGACGCCCGGGCTTTGC 77
Db 169 GTGAGCTTGGTGAGGCACCTTGAGCGGGCAACCTCGCCAGTATGAGCAGGAGTGCC 228
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTGCCCAAC 126
Db 229 CTGGGGCGGCACCATGAGAGCTCGAGCTCCTTGAGAGCGCTGCATCCGAC 277

RESULT 12
BZ410270
LOCUS
DEFINITION
OGAB01C ZM_0.7.1.5_KB Zea mays genomic clone ZMMEMa0020023,
genomic survey sequence.
ACCESSION
BZ410270
VERSION
BZ410270.1 GI:26042544
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 747)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source
1..747
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0020023"
/notes="Vector: pBCSK-; Site:1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 24.4%; Score 35.4; DB 28; Length 747;
Best Local Similarity 57.8%; Pred.No.1.9e+02; Indels 0; Gaps 0;
Matches 63; Conservative 0; Mismatches 46;
QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCCGACGCCCGGGCTTTGC 77
Db 593 GTGAGCTTGGTGAGGCACCTTGAGCGGGCAACCTCGCCAGTATGAGCAGGAGTGCC 534
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTGCCCAAC 126

```

```

Db      533 CTGGCGGCACCATTGAGAGCTCGAGCTCCTTGAGAGCCTGCATCCGAC 485

RESULT 13
CG285957/c
LOCUS
DEFINITION CG285957TH ZM.0.7.1.5_KB Zea mays genomic clone ZMMBMA0607P17,
genomic survey sequence.
ACCESSION CG285957
VERSION CG285957.1 GI:34200171
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 827)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGK93TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..827
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0607P17"
/clone_lib="ZM_0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Query Match 24.4%; Score 35.4; DB 29; Length 827;
Best Local Similarity 57.8%; Pred. No. 1.9e+02;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCCGACGCCCGCGCTTGC 77
Db 351 GTGAGCTTGTGAGGCACTTGCAGCGCGGCAACCTCCGCCAGTATGAGCAGGATGGC 292
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGGAGTGGCCAC 126
Db 291 CTGGCGGCACCATTGAGAGCTCGAGCTCCTTGAGAGCCTGCATCCGAC 243

RESULT 14
CG365468
LOCUS
DEFINITION CG2BF05TH ZM.0.7.1.5_KB Zea mays genomic clone ZMMBMA0751B10,
genomic survey sequence.
ACCESSION CG365468
VERSION CG365468.1 GI:34282735
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 839)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG3DB09TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0751B10"
/clone_lib="ZM_0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Query Match 24.4%; Score 35.4; DB 29; Length 839;
Best Local Similarity 57.8%; Pred. No. 1.9e+02;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCCGACGCCCGCGCTTGC 77
Db 244 GTGAGCTTGTGAGGCACTTGCAGCGCGGCAACCTCCGCCAGTATGAGCAGGATGGC 303
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGGAGTGGCCAC 126
Db 304 CTGGCGGCACCATTGAGAGCTCGAGCTCCTTGAGAGCCTGCATCCGAC 352

RESULT 15
CG441848
LOCUS
DEFINITION CG3DE09TH ZM.0.7.1.5_KB Zea mays genomic clone ZMMBMA079B18,
genomic survey sequence.
ACCESSION CG441848
VERSION CG441848.1 GI:34820478
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG3DB09TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA079B18"
/clone_lib="ZM_0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Query Match 24.4%; Score 35.4; DB 29; Length 839;
Best Local Similarity 57.8%; Pred. No. 1.9e+02;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCCGACGCCCGCGCTTGC 77
Db 244 GTGAGCTTGTGAGGCACTTGCAGCGCGGCAACCTCCGCCAGTATGAGCAGGATGGC 303
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGGAGTGGCCAC 126
Db 304 CTGGCGGCACCATTGAGAGCTCGAGCTCCTTGAGAGCCTGCATCCGAC 352

RESULT 15
CG441848
LOCUS
DEFINITION CG3DE09TH ZM.0.7.1.5_KB Zea mays genomic clone ZMMBMA079B18,
genomic survey sequence.
ACCESSION CG441848
VERSION CG441848.1 GI:34820478
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG3DB09TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA079B18"
/clone_lib="ZM_0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Query Match 24.4%; Score 35.4; DB 29; Length 839;
Best Local Similarity 57.8%; Pred. No. 1.9e+02;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCCGACGCCCGCGCTTGC 77
Db 244 GTGAGCTTGTGAGGCACTTGCAGCGCGGCAACCTCCGCCAGTATGAGCAGGATGGC 303
QY 78 CCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGGAGTGGCCAC 126
Db 304 CTGGCGGCACCATTGAGAGCTCGAGCTCCTTGAGAGCCTGCATCCGAC 352

```

methylation filtered genomic DNA library"

ORIGIN

Query Match	24.4%	Score 35.4;	DB 29;	Length 840;
Best Local Similarity	57.8%	Pred.No. 1.9e+02;		
Matches 63;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;

QY	18	GCGCGCTCGCTCGCTCACTAGGCGCGCGGACCAAGGTGCGCCGACGCCCGGCTTTC	77
Db	285	GTGAGCTTGGTGAGGCACTTGAGCGCGGCACTCCGCCAGTATGGAGCAGGATGGC	344

QY	78	CCGGCGGCGCTCAGTGAGCGAGCGCGGCGCAGAGAGGGAGTGGCCAC	126
Db	345	CTGGCGGCGCACCATTTGAGACTCGAGCTCCTTGAGAGCCTGCATCCGAC	393

Search completed: April 1, 2004, 17:02:14
 Job time : 1875.76 secs

RESULT 2

US-10-135-984-8
; Sequence 8, Application US/10135984
; Publication No. US20020182595A1
; GENERAL INFORMATION:
; APPLICANT: Matthew D. Weitzman
; APPLICANT: Anton J. Cathomen
; TITLE OF INVENTION: METHOD OF IDENTIFYING CELLULAR
; TITLE OF INVENTION: REGULATORS OF ADENO-ASSOCIATED VIRUS (AAV)
; FILE REFERENCE: SALKINS-041A
; CURRENT APPLICATION NUMBER: US/10/135,984
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 60/286951
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 146
; TYPE: DNA
; ORGANISM: adeno-associated virus
US-10-135-984-8

Query Match 100.0%; Score 145; DB 13; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGCGGCGAGCGGAGGAGTG 60
DB 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGCGGCGAGCGGAGGAGTG 60
QY 61 CGACGCCCGGCTTTCGCCGGGCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 120
DB 61 CGACGCCCGGCTTTCGCCGGGCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 3

US-09-782-378A-8
; Sequence 8, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-8

Query Match 100.0%; Score 145; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 60
DB 21 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 80
QY 61 CGACGCCCGGCTTTCGCCGGGCGGCTCAGTGAGCGGCGGCGAGCGGAGGAGTG 120

DB 81 CGACGCCCGGCTTTCGCCGGGCGGCTCAGTGAGCGGCGGCGAGCGGAGGAGTG 140
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 141 GCCAACTCCATCACTAGGGGTTCT 165

RESULT 4

US-10-054-665-7
; Sequence 7, Application US/10054665
; Publication No. US20020197237A1
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. US20020197237A1 5,478,745
US-10-054-665-7

Query Match 100.0%; Score 145; DB 13; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 60
DB 21 TTGGCCACTCCCTCTCTGGCGGCTCGCTCAGTGAGCGCGGCGGCGAGCGGAGGAGTG 80
QY 61 CGACGCCCGGCTTTCGCCGGGCGGCTCAGTGAGCGGCGGCGAGCGGAGGAGTG 120
DB 81 CGACGCCCGGCTTTCGCCGGGCGGCTCAGTGAGCGGCGGCGAGCGGAGGAGTG 140
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 141 GCCAACTCCATCACTAGGGGTTCT 165

RESULT 5

US-10-159-968-13/c
; Sequence 13, Application US/10159968
; Publication No. US20030152914A1
; GENERAL INFORMATION:
; APPLICANT: Musatov, Serge
; APPLICANT: Kaplitt, Michael G.
; TITLE OF INVENTION: Method for Generating Replication
; TITLE OF INVENTION: Defective Viral Vectors That are Helper Free
; FILE REFERENCE: 600-1-286
; CURRENT APPLICATION NUMBER: US/10/159,968
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/294,797
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: US 60/313,007
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Adeno-associated virus
US-10-159-968-13

Query Match 100.0%; Score 145; DB 14; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 60
DB 145 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 86

QY 61 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 120
DB 85 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 26

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 25 GCCAACTCCATCACTAGGGGTTCT 1

RESULT 6
US-10-276-356-1/c
; Sequence 1, Application US/10276356
; Publication No. US20040029106A1
; GENERAL INFORMATION:
; APPLICANT: University of No. US20040029106A1th Carolina at Chapel Hill
; APPLICANT: Samulski, R. Jude
; APPLICANT: McGarty, Douglas M.
; TITLE OF INVENTION: DUPLEXED PARVOVIRUS VECTORS
; FILE REFERENCE: 5470-282
; CURRENT APPLICATION NUMBER: US/10/276,356
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: PCT/US01/17587
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Inverted terminal repeat from the AAV-2 vector plasmid pSub 501
US-10-276-356-1

Query Match 100.0%; Score 145; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 60
DB 150 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 91

QY 61 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 120
DB 90 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 31

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 30 GCCAACTCCATCACTAGGGGTTCT 6

RESULT 7
US-10-023-208-58
; Sequence 58, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic enzyme attachment site sequence
US-10-023-208-58

Query Match 100.0%; Score 145; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.5e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 60
DB 42 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 101

QY 61 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 120
DB 102 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 161

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 162 GCCAACTCCATCACTAGGGGTTCT 186

RESULT 8
US-09-845-416-26
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match 100.0%; Score 145; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTGCGCC 60

QY 61 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCGCGGCGGCGCTCACTAGCGAGCGCGCAGAGAGGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 9
US-09-845-416-26/c
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142


```
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 145; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 60
Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 896

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 895 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 836

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 835 GCCAACTCCATCACTAGGGGTTCT 811

RESULT 10
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 145; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 60
Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 896

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 895 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 836

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 835 GCCAACTCCATCACTAGGGGTTCT 811

RESULT 11
US-09-845-416-33/c
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
```

```
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 145; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 60
Db 987 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 928

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 927 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 868

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 867 GCCAACTCCATCACTAGGGGTTCT 843

RESULT 12
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match      100.0%; Score 145; DB 10; Length 4414;
Best Local Similarity 100.0%; Pred. No. 9.8e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGGCGACCAAGGTCGCC 60

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 13
US-09-845-416-32/c
; Sequence 32, Application US/09845416
```

```
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 4414
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-32

Query Match      100.0%; Score 145; DB 10; Length 4414;
Best Local Similarity 100.0%; Pred. No. 9.8e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 60
DB 4414 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 4355
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 120
DB 4354 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 4295
QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 4294 GCCAACTCCATCACTAGGGTTTCCT 4270

RESULT 14
US-09-845-416-31
Sequence 31, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-31

Query Match      100.0%; Score 145; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 9.7e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 60
DB 4476 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 4417
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 120
DB 4416 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 4357
QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 4356 GCCAACTCCATCACTAGGGTTTCCT 4332

Search completed: April 1, 2004, 18:35:59
Job time : 229.574 secs

Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-31

Query Match      100.0%; Score 145; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 9.7e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 60
DB 4476 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC 4417
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 120
DB 4416 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGGCGGCGCGAGAGGGAGTG 4357
QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 4356 GCCAACTCCATCACTAGGGTTTCCT 4332

Search completed: April 1, 2004, 18:35:59
Job time : 229.574 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 12:58:58 ; Search time 210.185 Seconds

(without alignments)
25226.461 Million cell updates/sec

Title: US-10-620-039-1 COPY 1 125

Perfect score: 125
Sequence: 1 TTGGCCACTCCCTCTCTGG.....CGCAGAGGGGACTGGCCAA 125

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 secs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database :
1: Geneseq 257am04.*
N: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	125	100.0	144	8	ACF35877	AAV-2 rig
2	125	100.0	145	2	AAO41448	AAV2 Inve
3	125	100.0	145	2	AAO03385	Strict in
C 4	125	100.0	145	2	AAO63408	Inverted
	125	100.0	145	2	AAO34295	Adeno-ass
	125	100.0	145	6	ABA02990	Adeno-ass
5	125	100.0	145	6	ABA02990	Adeno-ass
6	125	100.0	145	6	ABSA69884	Human ade
7	125	100.0	145	6	ABSA69884	Human ade
C 8	125	100.0	145	7	ACCS8491	Adeno ass
	125	100.0	145	7	ACCS8491	Adeno ass
	125	100.0	145	8	ACF35876	AAV-2 lef
9	125	100.0	145	8	ACF35876	AAV-2 lef
10	125	100.0	146	7	ABX33568	Adeno-ass
11	125	100.0	165	2	AAQ66769	Double-D
12	125	100.0	165	2	AAO49462	Adeno ass
13	125	100.0	165	4	AAO03535	Inverted
14	125	100.0	165	6	ABSA69886	Human ade
C 15	125	100.0	165	6	ABSA69886	Human ade
	125	100.0	192	7	ABV77279	Nucleotid
	125	100.0	192	2	AAO03384	Right-han
16	125	100.0	192	2	AAO03384	Right-han
17	125	100.0	207	6	AAO44621	165 bp en
18	125	100.0	955	6	AAO37254	Adeno-ass
C 19	125	100.0	955	6	AAO37254	Adeno-ass
	125	100.0	955	6	AAO37254	Adeno-ass
	125	100.0	987	6	AAO37261	Adeno-ass
20	125	100.0	987	6	AAO37261	Adeno-ass
C 21	125	100.0	987	6	AAO37261	Adeno-ass
	125	100.0	4414	6	AAO37260	Adeno-ass
	125	100.0	4414	6	AAO37260	Adeno-ass
22	125	100.0	4414	6	AAO37260	Adeno-ass
C 23	125	100.0	4414	6	AAO37260	Adeno-ass
	125	100.0	4414	6	AAO37260	Adeno-ass

ALIGNMENTS

RESULT 1

ACE35877

ID ACF35877 standard; DNA; 144 BP.

XX

AC ACF35877;

XX XX

DT 06-NOV-20

XX XX

AAV-2 rig

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Adenovirus KW

1000

Adeno asso

0
2
3
9
7
6
5
4
3
2
1
0

0 X
2 X
3
9
7
6
5
4
3
2
1
0

PN WO2003061

XX
XX

31-III.-20

XX
XX
03-PRO-TC 93

17-TAN-20

CC	inverted terminal repeats. The vector is safe for use in gene therapy,
CC	partic. in treatment of haemoglobinopathies and a variety of diseases,
CC	e.g. thalassemia, diabetes, sickle cell anaemia, and cancer. See also
CC	AAQ41449. (Updated on 25-MAR-2003 to correct PN field.)
XX	
QQ	Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
	Query Match 100.0%; Score 125; DB 2; Length 145;
	Best Local Similarity 100.0%; Pred. No. 1.2e-23;
	Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 TTGGCAGCTCCCTCTCTGGCGCTCGCTCGCTCAGTGAGGCGGGCGACCAAGGTCGCC 60
Db	1 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCAGTGAGGCGGGCGACCAAGGTCGCC 60
Qy	61 CGAGCGCCGGCGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGGGAGTG 120
Db	61 CGAGCGCCGGCGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCAGAGGGGAGTG 120
Qy	121 GCCAA 125
Db	121 GCCAA 125
RESULT 3	
AA03385	
ID	AA03385 standard; cDNA; 145 BP.
AC	AA03385;
XX	
DT	16-OCT-2003 (revised)
DT	18-APR-1996 (first entry)
XX	
DE	Strict inverted terminal repeat from AAV-2, used in pITRFL.
XX	
KW	inverted terminal repeat; ITR, adenovirus; adeno-associated virus; AAV-2;
KW	replication defective; integration; gene therapy; ds.
XX	
OS	Adeno-associated virus 2.
XX	
PN	W09523867-A1.
XX	
PD	08-SEP-1995.
XX	
PF	28-FEB-1995; 95WO-FR000233.
XX	
PR	03-MAR-1994; 94FR-00002445.
XX	
PA	(RHON) RHONE POULENC RORER SA.
XX	
PI	Denefle P, Latta M, Ferricaudet M, Vigne E;
XX	
DR	WPI; 1995-320581/41.
XX	
PT	Recombinant defective adenovirus contg. integratable expression cassette
PT	- for use in gene therapy to express protein, antigen or anti:sense
PT	nucleic acid, also for prodn. of recombinant adeno-associated viruses.
XX	
PS	Example 3; Page 26; 50pp; French.
XX	
CC	Recombinant, non-pathogenic adenovirus which are able to integrate stably
CC	into a host genome are claimed. The viruses pref. contain at least one
CC	inverted terminal repeat (ITR) sequence and in particular two ITRs flank
CC	a heterologous DNA insert. The present sequence is that of the strict ITR
CC	from adeno associated virus AAV-2 (i.e. the ITR sequence without any
CC	deletions or additions). In the construct pITRFL, the beta-galactosidase
CC	marker gene is flanked by two strict AAV-2 ITRs. The defective viruses
CC	are useful for stably introducing large fragments of heterologous DNA
CC	making them suitable for gene therapy. (Updated on 16-OCT-2003 to
CC	standardise OS field)
XX	
QQ	Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 125; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCTCACTGAGCGCGGCGGCGACCAAGGTGCGCC 60
DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCTCACTGAGCGCGGCGGCGACCAAGGTGCGCC 60
QY 61 CGAGCGCGCGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGAGGAGGAGTG 120
DB 61 CGAGCGCGCGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGAGGAGGAGGAGTG 120
QY 121 GCCAA 125
DB 121 GCCAA 125

RESULT 4
AAT63408/c
ID AAT63408 standard; DNA; 145 BP.
AC AAT63408;
XX
XX
DT 18-DEC-1997 (first entry)
XX
DE Inverted terminal repeat from adeno associated virus serotype 2.
XX
XX Inverted terminal repeat; ITR; wild type; adeno associated virus; AAV;
KW serotype 2; Rep binding site; integration system; treatment; acquired;
KW inherited; disease; gene therapy; ss.
XX
OS Adeno associated virus.
XX
FH Key Location/Qualifiers
FT repeat_unit 1..42
FT /tag= a
FT /rpt_type= inverted
FT /note= "region A"
FT
FT misc_binding 20..31
FT /tag= g
FT /label= Rep_binding_site
FT repeat_unit 43..51
FT /tag= c
FT /rpt_type= inverted
FT /note= "region C"
FT
FT repeat_unit 53..61
FT /tag= d
FT /rpt_type= inverted
FT /note= "region C"
FT
FT repeat_unit 64..73
FT /tag= e
FT /rpt_type= inverted
FT /note= "region B"
FT
FT repeat_unit 75..84
FT /tag= f
FT /rpt_type= inverted
FT /note= "region B"
FT
FT repeat_unit 85..125
FT /tag= b
FT /rpt_type= inverted
FT /note= "region A"
FT
FT misc_feature 126..145
FT /tag= h
FT /note= "region D"
FT
FT
XX WO9709442-A1.
XX
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US014312.
XX
XX 08-SEP-1995; 95US-00525866.
XX
XX
```

```
(AVIG-) AVIGEN INC.
XX
XX Natsoulis G, Surosky RT;
XX
XX WPI; 1997-192917/17.
XX
XX Nucleic acid construct for integration into target site in recipient
PT genome - comprises target sequence to bind AAV Rep and heterologous
PT nucleotide sequence , for treatment of acquired or inherited disease.
XX
XX Claim 3; Fig 1; 54pp; English.
XX
XX The present sequence is the inverted terminal repeat (ITR) from the wild
CC type adeno associated virus (AAV) serotype 2 genome. A novel nucleic acid
CC construct, comprises a targetting sequence capable of binding AAV Rep,
CC e.g. the present sequence, and at least 1 heterologous nucleotide
CC sequence arranged relative to the targetting sequence, so that it is
CC integrated into a target site in a recipient genome. The integration
CC system can be used to integrate a selected nucleotide sequence into a
CC recipient cell genome, useful in the treatment of an acquired or
CC inherited disease, e.g. long term gene therapy. The integration system
CC has the site specific integration characteristics of AAV, is not limiting
CC in regard to the size of the nucleotide sequence to be delivered and does
CC not result in the integration of viral gene sequences into the recipient
CC cell genome. Specifically it can be used to safely and efficiently
CC deliver, and integrate large nucleotide sequences into a suitable
CC recipient cell genome without the risk of causing insertional mutagenesis
CC due to random integration events
XX
XX Sequence 145 BP; 20 A; 52 C; 49 G; 24 T; 0 U; 0 Other;
XX
XX Query Match      100.0%; Score 125; DB 2; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCTCACTGAGCGCGGCGGCGACCAAGGTGCGCC 60
DB 125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCTCACTGAGCGCGGCGGCGACCAAGGTGCGCC 66
QY 61 CGAGCGCGCGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGAGGAGGAGTG 120
DB 65 CGAGCGCGCGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGAGGAGGAGGAGTG 6
QY 121 GCCAA 125
DB 5 GCCAA 1

RESULT 5
AAX34295
ID AAX34295 standard; DNA; 145 BP.
XX
XX AAX34295;
XX
XX 16-JUL-1999 (first entry)
XX
XX Adeno-associated virus inverted terminal repeat sequence.
DE
XX
XX Recombinant; chimeric; parvovirus; adeno-associated virus; AAV; vector;
KW promoter; rep; cap; inverted terminal repeat; ITR; erythroid cell;
KW integration; gene expression; bone marrow; peripheral blood cell;
KW endothelial cell; myocardial cell; ss.
XX
XX Adeno-associated virus.
XX
XX WO9918227-A1.
XX
XX 15-APR-1999.
XX
XX 08-OCT-1998; 98WO-US021202.
XX
XX 08-OCT-1997; 97US-0061364P.
XX
XX
```

```

PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX
XX Srivastava A, Ponnazhagan S;
XX
XX WPI; 1999-264033/22.
XX
XX New recombinant adeno-associated vectors.
XX
XX Claim 2; Page 69; 76pp; English.
XX
XX The invention relates to new recombinant chimeric parvovirus-adeno-
CC associated virus (AAV) vectors comprising a promoter e.g. an AAV promoter
CC (AAK34296) and a selected DNA sequence, especially an AAV rep gene and a
CC parvovirus B19 cap gene, located between 2 AAV inverted terminal repeats
CC (ITR) such as the ITR sequence shown here. The system can specifically
CC target primitive progenitor and differentiated cells of the erythroid
CC lineage, and can achieve stable integration and expression of transduced
CC genes. The vectors can be used for the in vitro or in vivo delivery of
CC genes to cells such as bone marrow cells, peripheral blood cells,
CC endothelial cells and myocardial cells
XX
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 125; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTGC 60
QY 61 CGAGCGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGAGAGGGAGTG 120
DB 61 CGAGCGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGAGAGGGAGTG 120
QY 121 GCCAA 125
DB 121 GCCAA 125
XX
XX ABA02990 standard; DNA; 145 BP.
XX
XX ABA02990;
XX
XX 29-AUG-2003 (revised)
XX 19-FEB-2002 (first entry)
XX
XX Adeno-associated virus 2 ITR SEQ ID NO 2.
XX
XX Cytostatic; virucide; apoptosis inducer; p53; Saos-2 cell; infection;
XX cancer; virus; HPV16; HPV18; adeno-associated virus 2; AAV-2; ss.
XX
XX Adeno-associated virus 2.
XX
XX Key Location/Qualifiers
XX misc_structure 1..145
XX /tag= a
XX /label= ITR
XX misc_feature 72
XX /tag= b
XX /note= "unpaired base"
XX misc_feature 94
XX /tag= c
XX /note= "unpaired base"
XX
XX WO200180840-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-GE001795.
XX
XX

```

```

PR 20-APR-2000; 2000GB-00009887.
XX
XX (BTGI-) BTG INT LTD.
XX
XX Raj K, Beard PM;
XX
XX WPI; 2002-041365/05.
XX
XX Single stranded and/or looped DNA for treating mutant p53 associated
PT cancer or infection that inhibit cellular p53, having portion with an
PT base, internally located with respect to any 3' and 5' ends of the DNA.
XX
XX Example 11; Page 38; 51pp; English.
XX
XX The invention relates to single stranded and/or looped DNA having a
CC portion with at least one base, internally located with respect to any 3',
CC and 5' ends of the DNA, that is unbasepaired with another base in a form
CC that is capable of being internalised within a target cell, for use in
CC therapy, with cytostatic and virucide activity. The DNA acts as an
CC apoptosis inducer in cells that lack p53 functionality, useful for
CC killing a cell, preferably a dividing cell where the cell is other than a
CC Saos-2 cell and the DNA is not configured to express the peptide or
CC protein that selectively kills the cell. The DNA is preferably in the
CC form of an AAV or associated with AAV protein which has been treated such
CC that the DNA is no longer capable of replication or expression in cells
CC and is associated with or contained within a vehicle which is associated
CC with one or more viral fibers which facilitate internalisation of the DNA
CC into a target cell. The DNA is also useful for manufacturing a medicament
CC for treating an individual suffering from a mutant p53 associated cancer
CC or an infection that inhibits cellular p53. The method targets cancer
CC cells or cells infected with p53 inhibiting viruses, such as HPV16 or
CC HPV18 and only cells that lack p53 activity are killed and no damage to
CC cellular DNA is involved. The present sequence is that of the adeno-
CC associated virus (AAV) 2 ITR, the double loop structure found at each end
CC of the AAV DNA. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 145 BP; 24 A; 49 C; 52 G; 20 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTGC 60
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTGC 80
QY 61 CGAGCGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145
XX
XX RESULT 7
XX ABS69884
XX ID ABS69884 standard; DNA; 145 BP.
XX
XX ABS69884;
XX
XX 21-NOV-2002 (first entry)
XX
XX Human adeno-associated virus 2 terminal repeat sequence.
XX
XX Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
XX adenosine deaminase deficiency; severe combined immune deficiency; PAH;
XX beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
XX low density lipoprotein gene; familial hypercholesterolaemia;
XX hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
XX phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
XX dystrophin gene; muscular dystrophy; cystic fibrosis; immunosuppressant;
XX human cystic fibrosis transmembrane conductance regulator gene;
XX

```

KW antianemic; antilipemic; nootropic; cytostatic; dermatological;
 KW human adeno-associated virus 2; AAV2; terminal repeat; ds.
 XX Homo sapiens.
 OS US2002102731-A1.
 PN 01-AUG-2002.
 PD 12-FEB-2001; 2001US-00782378.
 PF 02-OCT-2000; 2000US-0237747P.
 PR (UNY) UNIV NEW YORK STATE RES FOUND.
 XX Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
 PI WPI; 2002-690619/74.
 DR Producing vector, by introducing vector having nucleotide sequence,
 PT adenoassociated virus inverted terminal repeats and packaging sequence, and adeno-
 PT associated virus terminal repeat, into cell, and culturing cell.
 XX Disclosure; Page 8; 191pp; English.
 PS The present invention relates to a new method of producing a vector. The
 CC method involves introducing recombinant vector having nucleotide sequence
 CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
 CC adenoassociated virus flanking NS, adenoassociated virus terminal repeat linked to inverted
 CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
 CC end of NS, into cell expressing adenoassociated virus early gene lacking from vector
 CC ; and culturing cell to produce another vector. The method is useful for
 CC generating vectors, especially mad vectors. The method is useful in
 CC transferring nucleotide sequences of interest into a cell, for gene
 CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
 CC The nucleotide sequences are useful for treating diseases associated with
 CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
 CC deficiency with severe combined immune deficiency, beta-chain of
 CC haemoglobin gene associated with beta-thalassaemia and sickle cell
 CC disease, receptor for low density lipoprotein gene associated with
 CC familial hypercholesterolaemia, hypoxanthine-guanine
 CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome.
 CC Phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
 CC dystrophin gene associated with muscular dystrophy, and human cystic
 CC fibrosis transmembrane conductance regulator gene associated with cystic
 CC fibrosis. The present nucleic acid sequence represents a human adeno-
 CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
 CC methods of the invention
 XX SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 125; DB 6; Length 145;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGCC 60
 Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGCC 60
 Qy 61 CGACGCCCGGGGTTTGGCGGGGCTTTCAGTGAGCGAGCGCGAGAGAGGAGTG 120
 Db 61 CGACGCCCGGGGTTTGGCGGGGCTTTCAGTGAGCGAGCGCGAGAGAGGAGTG 120
 Qy 121 GCCAA 125
 Db 121 GCCAA 125
 RESULT 8
 ACC58491/c
 ID ACC58491 standard; DNA; 145 BP.
 XX ACC58491;
 AC ACC58491;

XX 26-AUG-2003 (first entry)
 DT Adeno associated virus inverted terminal repeat region.
 XX AAV; inverted terminal repeat; insect; gene therapy; vector; ss.
 DE Adeno associated virus.
 XX Key Location/Qualifiers
 FH stem_loop 21..145
 FT /*tag= a
 FT XX
 PN WO2003042361-A2.
 XX 22-MAY-2003.
 PD 08-NOV-2002; 2002WO-US035829.
 XX 09-NOV-2001; 2001US-00986618.
 PR 13-AUG-2002; 2002US-00216870.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Kotin RM, Urabe M, Ding C;
 XX WPI; 2003-449568/42.
 DR Producing an adeno-associated virus (AAV) in an insect cell, e.g. for
 PT gene therapy, comprises introducing an insect cell-compatible vector into
 PT an insect cell and maintaining the insect cell under conditions that
 PT produce AAV.
 XX Disclosure; Fig 5A; 84pp; English.
 PS The present sequence is that of an adeno associated virus (AAV)
 CC palindromic inverted terminal repeat (ITR) region. An AAV ITR is used in
 CC novel insect cell-compatible vectors of the invention that are designed
 CC for the production of AAV in insect cells. The vectors comprise an AAV
 CC ITR and coding sequences for AAV VP1, VP2 and VP3 capsid proteins, for
 CC AAV non-structural replication (Rep) proteins Rep52 or Rep40, and for AAV
 CC Rep78 or Rep68. The method of producing AAV in insect cells provides an
 CC efficient, safe and economical means of producing a large amount of
 CC recombinant AAV particles which may be used in gene therapy
 XX SQ Sequence 145 BP; 23 A; 49 C; 52 G; 21 T; 0 U; 0 Other;
 Query Match 100.0%; Score 125; DB 7; Length 145;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGCC 60
 Db 145 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGCC 86
 Qy 61 CGACGCCCGGGGTTTGGCGGGGCTTTCAGTGAGCGAGCGCGAGAGAGGAGTG 120
 Db 85 CGACGCCCGGGGTTTGGCGGGGCTTTCAGTGAGCGAGCGCGAGAGAGGAGTG 26
 Qy 121 GCCAA 125
 Db 25 GCCAA 21
 RESULT 9
 ACF35876
 ID ACF35876 standard; DNA; 145 BP.
 XX ACF35876;
 XX ACF35876;
 DT 06-NOV-2003 (first entry)
 XX AAV-2 left inverted terminal repeat sequence.

```

XX Adenovirus; Rep78; Rep68; CAP; Gene therapy; AAV-2; ITR; ds.
XX Adeno associated virus.
XX WO2003061582-A2.
XX 31-JUL-2003.
XX 17-JAN-2003; 2003WO-US001624.
XX 18-JAN-2002; 2002US-0349532P.
XX (UYDU-) UNIV DUKE.
XX Li C, Zhang X;
XX WPI; 2003-627412/59.
XX New recombinant adenovirus comprising an adenovirus that encodes one or
XX more AAV REP78/68 polypeptides that are inducibly expressed, useful as a
XX vector for gene therapy.
XX Example; Page 120; 122pp; English.
XX The invention relates to a recombinant adenovirus comprising an
XX adenovirus that encodes one or more AAV REP78/68 polypeptides that are
XX inducibly expressed. A complete virus-mediated system for recombinant AAV
XX production is provided which comprises: (a) a first recombinant
XX adenovirus encoding one or more AAV REP78/68 polypeptides and one or more
XX viral capsid polypeptides; (b) a second recombinant adenovirus comprising
XX a gene of interest and AAV inverted terminal repeats that flank the gene
XX of interest; (c) viral helper functions; and (d) a host cell comprising
XX the first recombinant adenovirus, the second recombinant adenovirus, and
XX the viral helper functions. The recombinant AAV is useful as a vector for
XX gene therapy. The present sequence represents an AAV-2 left inverted
XX terminal repeat (ITR) sequence
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 8; Length 145;
Best Local Similarity 100.0%; Pred No 1,2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGTCCGC 60
QY 61 CGACGCCCGGGTTTCCCGCGCGCTCAGTGAGCGCGCGAGCGGAGGAGGAGTG 120
Db 61 CGACGCCCGGGTTTCCCGCGCGCTCAGTGAGCGCGCGAGCGGAGGAGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125
RESULT 10
ID ABX93568
XX ABX93568 standard; DNA; 146 BP.
XX AC ABX93568;
XX DT 23-MAY-2003 (first entry)
XX DE Adeno-associated virus-2 inverted terminal repeat.
XX KW Inverted terminal repeat; Rep recognition sequence; AAV-2; ss; RRS;
XX KM viral regulatory element.
XX OS Adeno-associated virus serotype 2.
XX FH Key Location/Qualifiers

```

```

FT misc_binding 1..41
FT FT /*tag= a
FT FT /bound moiety= "Nucleotides 125-85 of the present
FT FT sequence"
FT FT 42..50
FT FT /*tag= b
FT FT /bound moiety= "Nucleotides 62-54 of the present
FT FT sequence"
FT FT 54..62
FT FT /*tag= c
FT FT /bound moiety= "Nucleotides 50-42 of the present
FT FT sequence"
FT FT 64..72
FT FT /*tag= d
FT FT /bound moiety= "Nucleotides 84-76 of the present
FT FT sequence"
FT FT 76..84
FT FT /*tag= e
FT FT /bound moiety= "Nucleotides 72-64 of the present
FT FT sequence"
FT FT 85..125
FT FT /*tag= f
FT FT /bound moiety= "Nucleotides 41-1 of the present sequence"
FT FT 93..108
FT FT /*tag= g
FT FT /label= Rep recognition sequence
FT FT /note= "This RRS is specifically claim in claim 13"
FT FT 122..127
FT FT /*tag= h
FT FT /label= Terminal_resolution_site
FT FT
US2002182595-A1.
05-DEC-2002.
29-APR-2002; 2002US-00135984.
27-APR-2001; 2001US-0286951P.
(WEIT/) WEITZMAN M D.
(CATH/) CATHOMEN A J.
Weitzman MD, Cathomen AJ;
WPI; 2003-328607/31.
Identification of cellular regulations of adeno-associated virus involves
contacting an adeno-associated viral regulatory interacting molecule with
a nucleic acid comprising a adeno-associated viral regulatory element.
Disclosure; Fig 1; 16pp; English.
The invention relates to identifying molecules that interact with an
adeno-associated viral regulatory element (e.g. a Rep recognition
sequence, RRS) comprising contacting at least one molecule with a nucleic
acid comprising at least one adeno-associated viral regulatory element,
and identifying whether at least one molecule is bound to the viral
regulatory elements. Also included is a kit for analysing the interaction
between molecules and an adeno-associated viral regulatory element. The
method is useful for identifying molecules e.g. at least one cellular
protein, recombinant protein, synthetic protein or exogenous protein
(preferably human proteins that are derived from a cDNA library or
cellular lysate), or peptide, antibody, nucleic acid, lipid, carbohydrate
and/or organic or inorganic compound that interacts with an adeno-
associated viral regulatory element comprising an inverted terminal
repeat (ITR) or region. The method broadens the understanding of adeno-
associated virus (AAV) and facilitates the application of adeno-
associated virus-based vector system, therefore ITR-binding proteins and
other molecules provide new insights into the AAV life cycle, including
the regulation of gene expression and integration, and suggests
improvements for its application in therapy. The present sequence is an
AAV-2 ITR containing an RRS used in the method of the invention

```



```
SQ Sequence 146 BP; 21 A; 53 C; 49 G; 23 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 7; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCC 60
Dy 1 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCC 60

QY 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGTAGCGAGCGCGCAGAGGGAGTG 120
Dy 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGTAGCGAGCGCGCAGAGGGAGTG 120

QY 121 GCCAA 125
Dy 121 GCCAA 125

RESULT 11
AAQ66769
ID AAQ66769 standard; DNA; 165 BP.
XX
AC AAQ66769;
XX
DT 25-MAR-2003 (revised)
DT 20-JAN-1995 (first entry)
XX
D3 Double-D ITR terminal repeat.
XX
KW Double-D ITR; inverted terminal repeat; D-sequence; AAV;
KW adeno-associated virus; capsid; encapsidation; gene therapy; vector; ds.
XX
OS Synthetic.
XX
PN WO9413788-A1.
XX
PD 23-JUN-1994.
XX
PF 03-DEC-1993; 93WO-US011728.
XX
PR 04-DEC-1992; 92US-00989841.
XX
PA (UTPI-) UNIV PITTSBURGH.
XX
PI Samulski RJ, Xiao X;
XX
DR WPI; 1994-217868/26.
XX
PT Adenovirus associated viral inverted terminal repeat - for use in a
PT recombinant viral vector system for treatment of genetic diseases.
XX
PS Disclosure; Page 25; 44pp; English.
XX
CC The 20-bp D-sequence given in AAQ66773 is present in the inverted
CC terminal repeat (ITR) sequence of AAV and is required for viral
CC replication. A novel, modified terminal repeat structure, double-D ITR,
CC was constructed (AAQ66769) that contained a single 145 bp ITR sequence
CC with an additional B' sequence. The double-D ITR fragment allows
CC replication and encapsidation of recombinant DNA into AAV capsid
CC proteins. Vectors including double-D ITR can be used for gene replacement
CC therapies. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 165 BP; 27 A; 56 G; 26 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCC 60
Dy 21 TTGGCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCC 80
```

```
QY 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGTAGCGAGCGCGCAGAGGGAGTG 120
Dy 81 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGTAGCGAGCGCGCAGAGGGAGTG 140
QY 121 GCCAA 125
Dy 141 GCCAA 145

RESULT 12
AA49462
ID AA49462 standard; DNA; 165 BP.
XX
AC AA49462;
XX
DT 10-SEP-1997 (first entry)
XX
D3 Adeno associated virus inverted terminal repeat with double D region.
XX
KW Viral replication; RBP protein; inverted terminal repeat; ITR;
KW adeno-associated virus; AAV; vector; double-D; cis-acting;
KW lytic life cycle; gene therapy; ss.
XX
OS Adeno associated virus.
OS Synthetic.
XX
PH Key Location/Qualifiers
FT repeat_region 1..165
FT /tag= m
FT /label= Double-D
FT /note= "This 165 bp sequence, resulting from addition of
FT a second D repeat to the AAV ITR, has not been identified
FT in any naturally occurring virus"
FT repeat_unit 1..145
FT /tag= a
FT /rpt_type= INVERTED
FT /standard_name= "ITR"
FT /note= "This naturally occurring 145 bp ITR is located at
FT both ends of the AAV genome"
FT repeat_unit 1..20
FT /tag= b
FT /rpt_type= INVERTED
FT /label= D_repeat
FT misc_structure 21..145
FT /tag= 1
FT /label= T-shaped
FT /note= "The naturally occurring 145 bp ITR sequence can
FT form a T-shaped structure for DNA replication when single
FT -stranded"
FT repeat_unit 21..62
FT /tag= c
FT /rpt_type= INVERTED
FT /label= A' repeat
FT /note= "Forms the stem of a T-shaped structure when base
FT paired with repeat A"
FT stem_loop 61..82
FT /tag= j
FT /note= "Part of T-shaped structure, formed by base
FT pairing between the C' and C repeats"
FT repeat_unit 63..70
FT /tag= d
FT /rpt_type= INVERTED
FT /label= C' repeat
FT repeat_unit 74..81
FT /tag= e
FT /rpt_type= INVERTED
FT /label= C_repeat
FT stem_loop 84..104
FT /tag= k
FT /note= "Part of T-shaped structure, formed by base
FT pairing between the B' and B repeats"
FT repeat_unit 85..92
FT /tag= f
```

```

FT /rpt_type= INVERTED
FT /label= B_repeat
FT 96..103
FT *tag= g
FT /rpt_type= INVERTED
FT /label= B'_repeat
FT 104..145
FT *tag= n
FT /rpt_type= INVERTED
FT /label= A_repeat
FT /note= "Forms the stem of a T-shaped structure when base
FT paired with A' repeat"
FT 146..165
FT *tag= i
FT /label= D'
FT /note= "Additional D' sequence"
XX WO9636364-A1.
XX
XX 21-NOV-1996.
XX
XX 14-MAY-1996; 96WO-US06786.
XX
XX 15-MAY-1995; 95US-00440738.
XX
XX (SAMU/) SAMULSKI R J.
XX (XIAO/) XIAO X.
XX
XX Samulski RJ, Xiao X;
XX
XX WPI; 1997-042643/04.
XX
XX Double-D sequence directs adeno-associated virus integration into host
XX genome - used in gene therapy, maintains full length coding sequence of
XX therapeutic gene.
XX
XX Claim 1; Fig 9; 55pp; English.
XX
XX The adeno associated virus (AAV) has a 145 bp inverted terminal repeat
XX (ITR) located at each end of its genome. In addition to being able to
XX base pair with each other, the ITRs can also individually fold back on
XX themselves through the base pairing of A', B', B' and C, C' sequences
XX to form a T-shaped structure for DNA replication (see features table). It
XX has been found that viral mutants with deleted D sequences are unable to
XX replicate their DNA. Addition of a second D repeat (D') to the naturally,
XX occurring ITR resulted in a sequence, designated double-D, which was
XX sufficient to carry out the functions normally required of two wild-type
XX ITRs during a lytic AAV viral infection, i.e. it is capable of directing
XX replication and assembly into AAV, and/or the integration into the host
XX genome, of recombinant DNA containing the nucleic acid molecule. Vectors
XX and viral particles containing the double-D sequence are useful in gene
XX therapy. Replication and integration into the host genome is completely
XX effected through the double-D sequences, ensuring that the heterologous
XX gene sequences remain intact
XX
XX Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 125; DB 2; Length 165;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TTGGCCACTCCCTCTGCGCGTCTGCTCCTCACTGAGCGCGGCGACCAAGTTCGCC 60
DB 21 TTGGCCACTCCCTCTGCGCGTCTGCTCCTCACTGAGCGCGGCGACCAAGTTCGCC 80
OY 61 CGAGCGCGGCGCTTGGCCCGGGCGGCTCAGTGTAGCGGAGCGGCGGAGGAGGAGTG 120
DB 81 CGAGCGCGGCGCTTGGCCCGGGCGGCTCAGTGTAGCGGAGCGGCGGAGGAGGAGTG 140
OY 121 GCCAA 125
DB 141 GCCAA 145

```

```

RESULT 13
AAD03535
XX AAD03535 standard; DNA; 165 BP.
XX
XX AAD03535;
XX
XX 19-JUN-2001 (first entry)
XX
XX Inverted terminal repeat double DD DNA sequence.
XX
XX Recombinant adeno-associated virus; rAAV; circular intermediate; ITR;
XX inverted terminal repeat; haemostatic; antiscikling; neuroprotective;
XX antianaemic; nootropic; blood disorder; sickle cell anaemia;
XX thalassaemia; neurological disorder; haemophilia; Alzheimer's disease;
XX muscle disorder; Parkinson's disease; gene delivery; erythropoietin; epo;
XX CTR; cystic fibrosis transmembrane conductance receptor;
XX tyrosine hydroxylase; glucocerebrosidase gene; Gaucher's disease;
XX gene therapy; Double DD; ds.
XX
XX Unidentified.
XX
XX WO200125465-A1.
XX
XX 12-APR-2001.
XX
XX 06-OCT-2000; 2000WO-US027863.
XX
XX 07-OCT-1999; 99US-0158209P.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (ENGE/) ENGELHARDT J F.
XX (DONG/) DONGSHENG D.
XX (ZIYI/) ZIYING Y.
XX
XX Engelhardt JF, Dongsheng D, Ziyang Y;
XX
XX WPI; 2001-266321/27.
XX
XX Composition for transferring recombinant DNAs and to express a
XX polypeptide in a host cell, comprises two recombinant adeno-associated
XX viruses.
XX
XX Disclosure; Page 139; 144pp; English.
XX
XX The patent discloses a composition comprising at least two recombinant
XX adeno-associated viruses (rAAVs). Each rAAV contains a recombinant DNA
XX molecule comprising three DNA segments linked together. The first and the
XX third DNA segments comprise 5' and 3' inverted terminal repeats (ITRs)
XX respectively from the circular intermediate of AAV. The second DNA
XX segment in each virus is different and does not comprise AAV sequence. It
XX preferably comprises sequences encoding a therapeutically effective
XX polypeptide such as the cystic fibrosis transmembrane conductance
XX receptor gene (CFTR), the erythropoietin (epo) gene, the tyrosine
XX hydroxylase gene (Parkinson's syndrome) or the glucocere- brosidase gene
XX (Gaucher's disease). The circular intermediate of AAV imparts increased
XX episonal stability and persistence of the vector in the host cell.
XX Compositions comprising rAAV sequences are useful for transferring
XX recombinant DNAs to a host cell and express a polypeptide in a host cell.
XX The recombinant vector is useful in medical therapy, which includes
XX treatment or prophylaxis of blood disorders (e.g. sickle cell anaemia,
XX thalassaemia, haemophilia), neurological disorders, such as Alzheimer's
XX disease, Parkinson's disease, muscle disorders involving skeletal,
XX cardiac or smooth muscle. AAV vector is used as a delivery vehicle for
XX gene therapy. The present sequence is the ITR DD DNA sequence referred to
XX as "double sequence". This sequence is disclosed in U.S. Patent No.
XX 5,478,745. This sequence is not used in the present invention
XX
XX Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 125; DB 4; Length 165;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTGGCCACCTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGC 60
CC phenylalanine hydroxylase (PAH) gene associated with Iesch-Nyhan syndrome,
CC dystrophin gene associated with muscular dystrophy, and human cystic
CC fibrosis transmembrane conductance regulator gene associated with cystic
CC fibrosis. The present nucleic acid sequence represents a human adeno-
CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
CC methods of the invention
XX
SQ Sequence 165 BP; 24 A; 59 C; 53 G; 29 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACCTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGC 60
Db 21 TTGGCCACCTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGC 80
QY 61 CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGGACCAAGTCCGC 120
Db 81 CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGGACCAAGTCCGC 140
QY 121 GCCAA 125
Db 141 GCCAA 145
RESULT 14
ABS69886
ID ABS69886 standard; DNA; 165 BP.
XX
AC ABS69886;
XX
DT 21-NOV-2002 (first entry)
XX
DE Human adeno-associated virus 2 terminal repeat DD sequence.
XX
KW Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
KW adenosine deaminase deficiency; severe combined immune deficiency; PAH;
KW beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
KW low density lipoprotein gene; familial hypercholesterolaemia;
KW hypoxanthine-guanine phosphoribosyltransferase; Iesch-Nyhan syndrome;
KW phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
KW dystrophin gene; muscular dystrophy; cystic fibrosis; immunostimulant;
KW human cystic fibrosis transmembrane conductance regulator gene;
KW antianaemic; antilipaeamic; nootropic; cytostatic; dermatological;
KW human adeno-associated virus 2; AAV2; terminal repeat; ds.
XX
OS Homo sapiens.
XX
PN US2002102731-A1.
XX
PD 01-AUG-2002.
XX
PF 12-FEB-2001; 2001US-00782378.
XX
PR 02-OCT-2000; 2000US-0237747P.
XX
PA (UINY) UNIV NEW YORK STATE RES FOUND.
XX
PI Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
XX
DR WPI; 2002-690619/74.
XX
PT Producing vector, by introducing vector having nucleotide sequence,
PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
PT associated virus terminal repeat, into cell, and culturing cell.
XX
PS Disclosure; Page 8; 191pp; English.
XX
CC The present invention relates to a new method of producing a vector. The
CC method involves introducing recombinant vector having nucleotide sequence
CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
CC end of NS, into cell expressing adenovirus early gene lacking from vector
CC ; and culturing cell to produce another vector. The method is useful for
CC generating vectors, especially mad vectors. The method is useful in
CC transferring nucleotide sequences of interest into a cell, for gene
CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
CC The nucleotide sequences are useful for treating diseases associated with
CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
CC deficiency with severe combined immune deficiency, beta-chain of
CC haemoglobin gene associated with beta-thalassaemia and sickle cell
CC disease, receptor for low density lipoprotein gene associated with
CC familial hypercholesterolaemia, hypoxanthine-guanine

CC phosphoribosyltransferase associated with Iesch-Nyhan syndrome,
CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC dystrophin gene associated with muscular dystrophy, and human cystic
CC fibrosis transmembrane conductance regulator gene associated with cystic
CC fibrosis. The present nucleic acid sequence represents a human adeno-
CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
CC methods of the invention
XX
SQ Sequence 165 BP; 24 A; 59 C; 53 G; 29 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACCTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGC 60
Db 21 TTGGCCACCTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGC 80
QY 61 CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGGACCAAGTCCGC 120
Db 81 CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGGACCAAGTCCGC 140
QY 121 GCCAA 125
Db 141 GCCAA 145
RESULT 15
ABV77279/C
ID ABV77279 standard; DNA; 165 BP.
XX
AC ABV77279;
XX
DT 28-MAR-2003 (first entry)
XX
DE Nucleotide sequence of pPRT, a circular adeno-associated virus.
XX
KW Circular adeno-associated virus; cAAV; replication; hairpin;
KW gene therapy; pPRT; ss.
XX
OS Adeno associated virus.
XX
PN WO200297056-A2.
XX
PD 05-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017324.
XX
PR 31-MAY-2001; 2001US-0294797P.
PR 07-AUG-2001; 2001US-0313007P.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Kaplitt MG, Moussatov S;
XX
DR WPI; 2003-103706/09.
XX
PT Production of defective viral vectors for gene therapy that are
PT completely free of helper viral vectors and helper viruses.
XX
PS Example 4; Page 53; 69pp; English.
XX
CC The specification describes a nucleotide sequence capable of directing
CC circular adeno-associated virus (cAAV) replication. This nucleotide
CC sequence comprises a loop sequence (TGGCCAA) flanked on the 5' and 3'
CC sides by complementary sequences. The cAAVs are useful in gene
CC therapy, e.g. to treat an acute medical condition a nucleic acid encoding
CC a therapeutic protein is inserted into the cAAV. The present sequence
CC represents pPRT, a cAAV that contains a wild-type circularization point
CC (the PRT domain), consisting of a single ITR flanked by two D-sequences
XX
SQ Sequence 165 BP; 26 A; 56 C; 56 G; 27 T; 0 U; 0 Other;

Query Match 100.0%; Score 125; DB 7; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCCGC	60
DB	145	TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCCGC	86
QY	61	CGACGCGCGCGGCTTTGGCCCGCGCGCTCAGTGAGCGCGCGCGAGCGCGAGAGGAGTG	120
DB	85	CGACGCGCGCGGCTTTGGCCCGCGCGCTCAGTGAGCGCGCGCGAGCGCGAGAGGAGTG	26
QY	121	GCCAA 125	
DB	25	GCCAA 21	

Search completed: April 1, 2004, 15:07:23
Job time : 210.185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 16:04:14 ; Search time 188.426 Seconds
(without alignments)
2480.895 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125

Perfect score: 125
Sequence: 1 TTGGCCACTCCCTCTCGCG.....CGCAGAGGGAGTGGCCAA 125

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2465228 seqs, 1869859620 residues

Total number of hits satisfying chosen parameters: 4930456

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	130	9	US-09-928-158B-1
2	125	100.0	145	9	US-09-782-378A-6
3	125	100.0	145	14	US-10-240-198-2
4	125	100.0	146	13	US-10-135-984-8
5	125	100.0	165	9	US-09-782-378A-8
6	125	100.0	165	13	US-10-054-665-7
7	125	100.0	165	14	US-10-159-968-13
8	125	100.0	175	12	US-10-276-356-1
9	125	100.0	207	14	US-10-023-208-58
10	125	100.0	955	10	US-09-845-416-26
11	125	100.0	955	10	US-09-845-416-26
12	125	100.0	987	10	US-09-845-416-33
13	125	100.0	987	10	US-09-845-416-33
14	125	100.0	4414	10	US-09-845-416-32
15	125	100.0	4414	10	US-09-845-416-32

c 16	125	100.0	4476	10	US-09-845-416-31	Sequence 31, Appl
17	125	100.0	4476	10	US-09-845-416-31	Sequence 31, Appl
18	125	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
c 19	125	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
20	125	100.0	4675	9	US-09-782-378A-1	Sequence 1, Appl
21	125	100.0	4675	9	US-09-782-378A-2	Sequence 2, Appl
22	125	100.0	4675	14	US-10-240-198-2	Sequence 7, Appl
23	125	100.0	4675	14	US-10-291-583-7	Sequence 1, Appl
24	125	100.0	4675	9	US-09-804-898-1	Sequence 1, Appl
25	125	100.0	4679	9	US-09-945-681-10	Sequence 10, Appl
26	125	100.0	4679	13	US-10-038-972A-12	Sequence 12, Appl
27	125	100.0	4679	14	US-10-136-819-6	Sequence 6, Appl
28	125	100.0	4680	13	US-10-077-294-1	Sequence 1, Appl
29	125	100.0	4680	13	US-10-163-886-1	Sequence 1, Appl
30	125	100.0	4680	14	US-10-263-127-1	Sequence 1, Appl
31	125	100.0	4680	14	US-10-375-777-1	Sequence 1, Appl
32	125	100.0	4681	12	US-10-696-261-18	Sequence 18, Appl
33	125	100.0	4681	12	US-10-696-282-18	Sequence 18, Appl
34	125	100.0	4681	12	US-10-696-900-18	Sequence 18, Appl
35	125	100.0	4683	12	US-10-696-261-19	Sequence 19, Appl
36	125	100.0	4683	12	US-10-696-282-19	Sequence 19, Appl
37	125	100.0	4683	12	US-10-696-900-19	Sequence 19, Appl
38	125	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
c 39	125	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
40	125	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
c 41	125	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
c 42	125	100.0	4866	10	US-09-845-416-28	Sequence 28, Appl
c 43	125	100.0	4990	10	US-09-845-416-34	Sequence 34, Appl
c 44	125	100.0	4990	10	US-09-845-416-34	Sequence 34, Appl
45	125	100.0	5060	10	US-09-845-416-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-09-928-158B-1
; Sequence 1, Application US/09928158B
; Patent No. US30020177222A1
; GENERAL INFORMATION:
; APPLICANT: SIKUN, LI
; TITLE OF INVENTION: REPLICATION COMPETENT AAV HELPER FUNCTIONS
; FILE REFERENCE: 102182-18
; CURRENT APPLICATION NUMBER: US/09/928,158B
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/224,132
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 130
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
US-09-928-158B-1

Query Match	100.0%	Score 125;	DB 9;	Length 130;
Best Local Similarity	100.0%	Pred. No. 1.4e-25;		
Matches 125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTGCGC	60	
Db	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTGCGC	60	
QY	61	CGACGCCCGGGCTTTGCGCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG	120	
Db	61	CGACGCCCGGGCTTTGCGCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG	120	
QY	121	GCCTAA 125		
Db	121	GCCTAA 125		
RESULT 2				

```
US-09-782-378A-6
; Sequence 6, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-6

Query Match      100.0%; Score 125; DB 9; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 120
QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 4
US-10-135-984-8
; Sequence 8, Application US/10135984
; Publication No. US20020182595A1
; GENERAL INFORMATION:
; APPLICANT: Matthew D. Weitzman
; APPLICANT: Anton J. Cathomen
; TITLE OF INVENTION: METHOD OF IDENTIFYING CELLULAR
; TITLE OF INVENTION: REGULATORS OF ADENO-ASSOCIATED VIRUS (AAV)
; FILE REFERENCE: SALKINS.041A
; CURRENT APPLICATION NUMBER: US/10/135,984
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 60/286951
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 146
; TYPE: DNA
; ORGANISM: adeno-associated virus
US-10-135-984-8

Query Match      100.0%; Score 125; DB 13; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 5
US-09-782-378A-8
; Sequence 8, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 145
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
; FEATURE:
; NAME/KEY: misc structure
; LOCATION: (1)..(145)
; OTHER INFORMATION: ITR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (72)_feature
; OTHER INFORMATION: Unpaired base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (94)_feature
; OTHER INFORMATION: Unpaired base
US-10-240-198-2
; Sequence 2, Application US/10240198
; Publication No. US20030100115A1
; GENERAL INFORMATION:
; APPLICANT: BTG International Ltd
; APPLICANT: BEARD DR, PETER
; APPLICANT: RAJ DR, KENNETH
; TITLE OF INVENTION: CYTOTOXIC AGENTS
; FILE REFERENCE: 142184WO
; CURRENT APPLICATION NUMBER: US/10/240,198
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 0009887.1
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 145
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
; FEATURE:
; NAME/KEY: misc structure
; LOCATION: (1)..(145)
; OTHER INFORMATION: ITR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (72)_feature
; OTHER INFORMATION: Unpaired base
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (94)_feature
; OTHER INFORMATION: Unpaired base
US-10-240-198-2
```

```
; SEQ ID NO 8
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-8

Query Match      100.0%; Score 125; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGCGGCGACCAAAAGTGGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGCGGCGACCAAAAGTGGCC 80

QY 61 CGAGCGCCGGGCTTTGCCCGCGGCGCTCACTAGCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 81 CGAGCGCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 140

QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 6
US-10-054-665-7
; Sequence 7, Application US/10054665
; Publication No. US20020197237A1
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665
; CURRENT FILING DATE: 2002-06-13
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. US20020197237A1 5,478,745
US-10-054-665-7

Query Match      100.0%; Score 125; DB 13; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 80

QY 61 CGAGCGCCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 81 CGAGCGCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCGCAGAGAGGAGTG 140

QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 7
US-10-159-968-13/C
; Sequence 13, Application US/10159968
; Publication No. US20030152914A1
; GENERAL INFORMATION:
; APPLICANT: Kaplitt, Michael G.
; APPLICANT: Musatov, Serge
; TITLE OF INVENTION: Method for Generating Replication
```

```
; TITLE OF INVENTION: Defective Viral Vectors That are Helper Free
; FILE REFERENCE: 600-1-286
; CURRENT APPLICATION NUMBER: US/10/159,968
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/294,797
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: US 60/313,007
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Adeno-associated virus
US-10-159-968-13

Query Match      100.0%; Score 125; DB 14; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 60
Db 145 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 86

QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 85 CGACGCCCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 26

QY 121 GCCAA 125
Db 25 GCCAA 21

RESULT 8
US-10-276-356-1/C
; Sequence 1, Application US/10276356
; Publication No. US20040029106A1
; GENERAL INFORMATION:
; APPLICANT: University of No. US20040029106A1h Carolina at Chapel Hill
; APPLICANT: Samulski, R. Jude
; APPLICANT: McCarty, Douglas M.
; TITLE OF INVENTION: DUPLEXED PARVOVIRUS VECTORS
; FILE REFERENCE: 5470-282
; CURRENT APPLICATION NUMBER: US/10/276,356
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: PCT/US01/17587
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Inverted terminal repeat from the AAV-2 vector plasmid pSub 201
US-10-276-356-1

Query Match      100.0%; Score 125; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 60
Db 150 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTAGCGCGGCGGCGACCAAAAGTGGCC 91

QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 90 CGACGCCCGGGCTTTGCCCGGGCGGCTCACTAGCGAGCGAGCGCGCGCAGAGAGGAGTG 31

QY 121 GCCAA 125
Db 30 GCCAA 26
```

```
RESULT 9
US-10-023-208-58
; Sequence 58, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Chiang
; TITLE OF INVENTION: PROCAROTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/AFT/RMS/RVK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic enzyme attachment site sequence
US-10-023-208-58

Query Match      100.0%; Score 125; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 60
DB 42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 101
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 120
DB 102 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 161
QY 121 GCCAA 125
DB 162 GCCAA 166

RESULT 10
US-09-845-416-26
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 125; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 60
DB 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 896
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 120
DB 895 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 836
QY 121 GCCAA 125
DB 835 GCCAA 831

RESULT 11
US-09-845-416-26/c
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 125; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 60
DB 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 896
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 120
DB 895 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 836
QY 121 GCCAA 125
DB 835 GCCAA 831

RESULT 12
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 125; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 60
DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGCGAGAGGAGTG 836
```


Qy 61 CGACGCCCGGGCTTTGCCCGCGGCCTCAGTGAAGCAGCGAGCGCGCAGAGGGAGTG 120
 |||||
 Db 61 CGACGCCCGGGCTTTGCCCGCGGCCTCAGTGAAGCAGCGAGCGCGCAGAGGGAGTG 120
 |||||
 Qy 121 GCCAA 125
 |||||
 Db 121 GCCAA 125
 |||||

RESULT 13

```

US-09-845-416-33/c
; Sequence 33, Application US/09845416
; Publication, No US20030171312A1
; GENERAL INFORMATION
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

```

RESULT 14

```

US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication NO. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

```

Db	1	TTGGCCACCTCCCTCTGTGGCGCTCGCTCAGTGAGCCCGGGCGACCAAGGTGCC	60
Qy	61	CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCACAGAGGGAGTG	120
Db	61	CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCACAGAGGGAGTG	120
Qy	121	GCCAA 125	
Db	121	GCCAA 125	

RESULT 15

```

US-09-845-416-32/C
; Sequence 32, Application US/09845416
; Publication No. US2003017312A1
; GENERAL INFORMATION:
; APPLICANT: XIACI, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; THEREOF
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

```

Search completed: April 1, 2004, 18:36:00
Job time : 189.426 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:52:43 ; Search time 61.7593 Seconds
(without alignments)
1302.928 Million cell updates/sec

Title: US-10-620-039-1

Perfect score: 145

Sequence: 1 TTGGCCATCTCTCTCTGGG.....CTCCATCACTAGGGTTCCT 145

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	145	1	US-07-789-917A-1
2	145	100.0	145	3	US-08-702-573-4
3	145	100.0	145	3	US-07-982-193-1
4	145	100.0	145	1	US-07-989-841A-1
5	145	100.0	145	2	US-08-440-738A-1
6	145	100.0	145	3	US-08-471-914-1
7	145	100.0	145	4	US-09-276-625-7
8	145	100.0	4680	1	US-08-254-358-1
9	145	100.0	4680	1	US-08-475-391-1
10	145	100.0	4680	2	US-08-709-609-1
11	145	100.0	4680	5	PCT-US95-07178-1
12	145	100.0	5932	4	US-09-299-141-4
13	145	100.0	5932	4	US-09-299-141-4
14	145	100.0	6142	4	US-09-299-141-8
15	145	100.0	6142	4	US-09-299-141-8
16	145	100.0	6253	3	US-08-893-327-15
17	145	100.0	6253	3	US-08-893-327-15
18	145	100.0	6280	3	US-08-893-327-17
19	145	100.0	6280	3	US-08-893-327-17
20	145	100.0	6280	3	US-08-893-327-19
21	145	100.0	6280	3	US-08-893-327-19
22	145	100.0	6565	4	US-09-299-141-1
23	145	100.0	6565	4	US-09-299-141-1
24	145	100.0	6714	4	US-09-299-141-6
25	145	100.0	6714	4	US-09-299-141-6
26	145	100.0	6924	4	US-09-299-141-9
27	145	100.0	6924	4	US-09-299-141-9

28 145 100.0 6924 4 US-09-299-141-10 Sequence 10, Appl
29 145 100.0 6924 4 US-09-299-141-10 Sequence 10, Appl
30 145 100.0 6924 4 US-09-299-141-11 Sequence 11, Appl
31 145 100.0 6924 4 US-09-299-141-11 Sequence 11, Appl
32 145 100.0 6981 4 US-09-299-141-7 Sequence 7, Appl
33 145 100.0 6981 4 US-09-299-141-7 Sequence 7, Appl
34 145 100.0 7054 4 US-09-299-141-3 Sequence 3, Appl
35 145 100.0 7054 4 US-09-299-141-3 Sequence 3, Appl
36 145 100.0 7405 4 US-09-299-141-2 Sequence 2, Appl
37 145 100.0 7405 4 US-09-299-141-2 Sequence 2, Appl
38 145 100.0 7492 4 US-09-299-141-5 Sequence 5, Appl
39 145 100.0 7492 4 US-09-299-141-5 Sequence 5, Appl
40 145 100.0 8698 4 US-09-770-315-2 Sequence 2, Appl
41 143.4 98.9 272 4 US-09-276-625-4 Sequence 4, Appl
42 141.8 97.8 5585 2 US-08-305-221-1 Sequence 1, Appl
43 141.8 97.8 5585 4 US-09-000-003A-1 Sequence 1, Appl
44 138.6 95.6 272 4 US-09-276-625-6 Sequence 6, Appl
45 135.4 93.4 145 1 US-07-989-841A-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-07-789-917A-1
; Sequence 1, Application US/07789917A
; Patent No. 5252479
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release q.0, Version q.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789, 917A
; FILING DATE: 19911118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-789-917A-1

Query Match 100.0%; Score 145; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCATCTCTCTCTGGGCTCGCTCACTAGGCGCGGCGACCAAGGTGCGC 60
Db 1 TTGGCCATCTCTCTCTGGGCTCGCTCACTAGGCGCGGCGACCAAGGTGCGC 60
QY 61 CGAGCGCGGCGCTTTTGGCGGCGCGCTCACTAGGCGGCGAGCGCGACAGAGGGAGTG 120

Db 61 CGACGCCCGGCTTTCCCGCGGGCTCAGTGAGCGAGCGCGGAGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 2
US-08-702-573-4
; Sequence 4, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTA, Martine
; APPLICANT: DENEPIE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: PERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3rd
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..145
; OTHER INFORMATION: /note= "Minimal ITR Sequence"
US-08-702-573-4

Query Match 100.0%; Score 145; DB 3; Length 145;
Best Local Similarity 100.0%; P-red. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGCTTTGCGCGGCTCGCTCACTGAGCGAGCGAGCGGAGGAGTG 120
Db 61 CGACGCCCGGCTTTGCGCGGCTCGCTCACTGAGCGAGCGGAGGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 3
US-07-982-193-1
; Sequence 1, Application US/07982193
; Patent No. 6261834
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,193
; FILING DATE: 19921125
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-982-193-1
Query Match 100.0%; Score 145; DB 3; Length 145;
Best Local Similarity 100.0%; P-red. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGCTTTGCGCGGCTCGCTCACTGAGCGAGCGGAGGAGGAGTG 120
Db 61 CGACGCCCGGCTTTGCGCGGCTCGCTCACTGAGCGAGCGGAGGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 4
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-013
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-07-989-841A-1
;
; Query Match 100.0%; Score 145; DB 1; Length 165;
; Best Local Similarity 100.0%; Pred. No. 4.3e-31;
; Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 60
; DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 80
;
; QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
; DB 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 140
;
; QY 121 GCCAACTCCATCACTAGGGTTCTT 145
; DB 141 GCCAACTCCATCACTAGGGTTCTT 165
;
; RESULT 5
; US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A
;
; Query Match 100.0%; Score 145; DB 1; Length 165;
; Best Local Similarity 100.0%; Pred. No. 4.3e-31;
; Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 60
; DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 80
;
; QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
; DB 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 140
;
; QY 121 GCCAACTCCATCACTAGGGTTCTT 145
; DB 141 GCCAACTCCATCACTAGGGTTCTT 165
;
; RESULT 6
; US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samulski, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; FILE REFERENCE: 6636-027
; CURRENT APPLICATION NUMBER: US/08/471,914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440,738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: double-D
; OTHER INFORMATION: sequence
; US-08-471-914-1
;
; Query Match 100.0%; Score 145; DB 3; Length 165;
; Best Local Similarity 100.0%; Pred. No. 4.3e-31;
; Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 60
; DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTCGCC 80
;
; QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
; DB 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 140
;
; QY 121 GCCAACTCCATCACTAGGGTTCTT 145
```

Db 141 GCCAACTCCATCACTAGGGGTTCT 165

RESULT 7
US-09-276-625-7
Sequence 7, Application US/09276625
Patent No. 6436392
GENERAL INFORMATION:
APPLICANT: Engelhardt, John F.
APPLICANT: Duan, Dongsheng
TITLE OF INVENTION: Adeno-associated virus vectors
FILE REFERENCE: 875.007U51
CURRENT APPLICATION NUMBER: US/09/276,625
CURRENT FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/086,166
PRIOR FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 165
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7

Query Match 100.0%; Score 145; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 80
QY 61 CGAGCGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGAGCGAGCGCGGAGGGAGTG 120
Db 81 CGAGCGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGAGCGAGCGCGGAGGGAGTG 140
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 141 GCCAACTCCATCACTAGGGGTTCT 165

RESULT 8
US-08-254-358-1
Sequence 1, Application US/08254358
Patent No. 5658785
GENERAL INFORMATION:
APPLICANT: Johnson, Philip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF INVENTION: 3
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,358
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5658785and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-475-391-1

Query Match 100.0%; Score 145; DB 1; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGGCGACCAAAAGTCCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGGCGACCAAAAGTCCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGCGGCGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGCGGCGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 10

US-08-709-609-1
Sequence 1, Application US/08709609
Patent No. 5858775
GENERAL INFORMATION:
APPLICANT: Johnson, Philip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,609
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5858775and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-709-609-1

Query Match 100.0%; Score 145; DB 2; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGGCGGCGGCGGCGGAGTG 60
DB 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGGCGGCGGCGGCGGAGTG 60
QY 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGCGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGCGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 11

PCT-US95-07178-1
Sequence 1, Application PC/TUS9507178
GENERAL INFORMATION:
APPLICANT: Johnson, Philip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07178
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

Query Match 100.0%; Score 145; DB 5; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGGCGGCGGCGGAGTG 60
DB 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGGCGGCGGCGGAGTG 60
QY 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGGCGGCGCTCACTGAGCGGCGGCGGCGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 12

US-09-299-141-4
Sequence 4, Application US/09299141
Patent No. 6461606
GENERAL INFORMATION:
APPLICANT: FLOTTE, TERENCE R.
APPLICANT: SONG, SIHONG
APPLICANT: BYRNE, BARRY J.
APPLICANT: MORGAN, MICHAEL
TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY

Qy	1	TTGGCCACTCCCTCTCTGTGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC	60
Db	3288	TTGGCCACTCCCTCTCTGTGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC	3229
Qy	61	CGACGCCCGGGCTTTGCCCCGGGGCTCAGTGAGCGAGCGCGCAGAGAGGGAGTG	120
Db	3228	CGACGCCCGGGCTTTGCCCCGGGGCTCAGTGAGCGAGCGCGCAGAGAGGGAGTG	3169
Qy	121	GCCAACTCCATCACTAGGGTTCTT	145
Db	3168	GCCAACTCCATCACTAGGGTTCTT	3144

Search completed: April 1, 2004, 17:04:17
Job time : 65.7593 secs

XX The sequence is one of two inverted terminal repeat sequences, one from
 CC parvovirus B19 and the other from adeno-associated virus 2 (AAV2), used
 CC in construction of an expression vector for site specific integration and
 CC cell specific gene expression. The vector comprises at least one cassette
 CC contg. a promoter capable of effecting cell-specific expression, operably
 CC linked to a hetero- logous gene, and the cassette residing between the
 CC inverted terminal repeats. The vector is safe for use in gene therapy,
 CC partic. in treatment of haemoglobinopathies and a variety of diseases,
 CC e.g. thalassemia, diabetes, sickle cell anaemia, and cancer. See also
 CC AAO41449. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;

Query Match 100.0%; Score 145; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
 DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
 QY 61 CGACGCCCGGGGTTTGGCCGGCGGCGCTCAGTACGAGCGAGCGCGCAGAGAGGAGTG 120
 DB 61 CGACGCCCGGGGTTTGGCCGGCGGCGCTCAGTACGAGCGAGCGCGCAGAGAGGAGTG 120
 QY 121 GCCAACTCCATCCTAGGGGTTTCCT 145
 DB 121 GCCAACTCCATCCTAGGGGTTTCCT 145

RESULT 2
 AAT03385
 ID AAT03385 standard; cDNA; 145 BP.
 XX
 AC AAT03385;
 XX
 DT 16-OCT-2003 (revised)
 DT 18-APR-1996 (first entry)
 XX
 DE Strict inverted terminal repeat from AAV-2, used in pITRFL.
 DE
 KW inverted terminal repeat; ITR; adenovirus; adeno-associated virus; AAV-2;
 KW replication defective; integration; gene therapy; ds.
 XX

OS Adeno-associated virus 2.
 XX
 PN WO9523867-A1.
 XX
 PD 08-SEP-1995.
 XX
 PF 28-FEB-1995; 95WO-FR000233.
 XX
 PR 03-MAR-1994; 94FR-00002445.
 XX
 PA (RHON) RHONE POULENC RORER SA.
 XX
 PI Denefle P, Latta M, Perricaudet M, Vigne E;
 XX
 DR WPI; 1995-320581/41.
 XX
 PT Recombinant defective adenovirus contg. integratable expression cassette
 PT - for use in gene therapy to express protein, antigen or anti-sense
 PT nucleic acid, also for prodn. of recombinant adeno-associated viruses.
 XX
 PS Example 3; Page 26; 50pp; French.
 XX

CC Recombinant, non-pathogenic adenovirus which are able to integrate stably
 CC into a host genome are claimed. The viruses pref. contain at least one
 CC inverted terminal repeat (ITR) sequence and in particular two ITRs flank
 CC a heterologous DNA insert. The present sequence is that of the strict ITR
 CC from adeno associated virus AAV-2 (i.e. the ITR sequence without any
 CC deletions or additions). In the construct pITRFL, the beta-galactosidase

CC marker gene is flanked by two strict AAV-2 ITRs. The defective viruses
 CC are useful for stably introducing large fragments of heterologous DNA
 CC making them suitable for gene therapy. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
 DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
 QY 61 CGACGCCCGGGGTTTGGCCGGCGGCGCTCAGTACGAGCGAGCGCGCAGAGAGGAGTG 120
 DB 61 CGACGCCCGGGGTTTGGCCGGCGGCGCTCAGTACGAGCGAGCGCGCAGAGAGGAGTG 120
 QY 121 GCCAACTCCATCCTAGGGGTTTCCT 145
 DB 121 GCCAACTCCATCCTAGGGGTTTCCT 145

RESULT 3
 AAX34295
 ID AAX34295 standard; DNA; 145 BP.
 XX
 AC AAX34295;
 XX
 DT 16-JUL-1999 (first entry)
 XX
 DE Adeno-associated virus inverted terminal repeat sequence.
 XX
 KW Recombinant; chimeric; parvovirus; adeno-associated virus; AAV; vector;
 KW promoter; rep; cap; inverted terminal repeat; ITR; erythroid cell;
 KW integration; gene expression; bone marrow; peripheral blood cell;
 KW endothelial cell; myocardial cell; ss.
 XX

OS Adeno-associated virus.
 XX
 PN WO9918227-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 08-OCT-1998; 98WO-US021202.
 XX
 PR 08-OCT-1997; 97US-0061364P.
 XX
 PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
 XX
 PI Srivastava A, Ponnazhagan S;
 XX
 DR WPI; 1999-264033/22.
 XX
 PT New recombinant adeno-associated vectors.
 XX
 PS Claim 2; Page 69; 76pp; English.
 XX

CC The invention relates to new recombinant chimeric parvovirus-adeno-
 CC associated virus (AAV) vectors comprising a promoter e.g. an AAV promoter
 CC (AAX34296) and a selected DNA sequence, especially an AAV rep gene and a
 CC parvovirus B19 cap gene, located between 2 AAV inverted terminal repeats
 CC (ITR) such as the ITR sequence shown here. The system can specifically
 CC target primitive progenitor and differentiated cells of the erythroid
 CC lineage and can achieve stable integration and expression of transduced
 CC genes. The vectors can be used for the in vitro or in vivo delivery of
 CC genes to cells such as bone marrow cells, peripheral blood cells,
 CC endothelial cells and myocardial cells

SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 2; Length 145;

Best Local Similarity 100.0%; Pred. No. 8.1e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60

QY 61 CGACGCCCGGGCTTTGGCCGGCGGCTCGCTCACTGAGCGGCGGCGACGAGGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGGCCGGCGGCTCGCTCACTGAGCGGCGGCGACGAGGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 121 GCCAACTCCATCACTAGGGTTTCCT 145

RESULT 4
ABS69884
ID ABS69884 standard; DNA; 145 BP.

XX ABS69884;
XX
XX
XX 21-NOV-2002 (first entry)
XX
XX
XX Human adeno-associated virus 2 terminal repeat sequence.

XX Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
XX adenosine deaminase deficiency; severe combined immune deficiency; PAH;
XX beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
XX low density lipoprotein gene; familial hypercholesterolaemia;
XX hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
XX phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
XX dystrophin gene; muscular dystrophy; cystic fibrosis; immunostimulant;
XX human cystic fibrosis transmembrane conductance regulator gene;
XX antianemic; antileptemic; nototropic; cytostatic; dermatological;
XX human adeno-associated virus 2; AAV2; terminal repeat; ds.

XX Homo sapiens.
XX
XX
XX US2002102731-A1.
XX
XX 01-AUG-2002.
XX
XX 12-FEB-2001; 2001US-00782378.
XX
XX 02-OCT-2000; 2000US-0237747P.
XX
XX (UYN) UNIV NEW YORK STATE RES FOUND.
XX
XX Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
XX
XX WPI; 2002-690619/74.
XX
XX Producing vector, by introducing vector having nucleotide sequence,
XX adenovirus inverted terminal repeats and packaging sequence, and adeno-
XX associated virus terminal repeat, into cell, and culturing cell.
XX
XX Disclosure; Page 8; 191pp; English.

XX The present invention relates to a new method of producing a vector. The
XX method involves introducing recombinant vector having nucleotide sequence
XX (NS) having 5' and 3' end, left and right inverted terminal repeats of
XX adenovirus flanking NS, adenovirus packaging sequence linked to inverted
XX terminal repeat, and adeno-associated virus terminal repeat linked to 3'
XX end of NS, into cell expressing adenovirus early gene lacking from vector
XX ; and culturing cell to produce another vector. The method is useful for
XX generating vectors, especially mAd vectors. The method is useful in
XX transferring nucleotide sequences of interest into a cell, for gene
XX transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
XX The nucleotide sequences are useful for treating diseases associated with
XX it, i.e. adenosine deaminase gene associated with adenosine deaminase
XX deficiency with severe combined immune deficiency, beta-chain of
XX haemoglobin gene associated with beta-thalassaemia and sickle cell

CC disease, receptor for low density lipoprotein gene associated with
CC familial hypercholesterolaemia, hypoxanthine-guanine
CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC dystrophin gene associated with muscular dystrophy, and human cystic
CC fibrosis transmembrane conductance regulator gene associated with cystic
CC fibrosis. The present nucleic acid sequence represents a human adeno-
CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
CC methods of the invention
XX
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 145; DB 6; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 8.1e-30;
XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60

QY 61 CGACGCCCGGGCTTTGGCCGGCGGCTCGCTCACTGAGCGGCGGCGACGAGGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGGCCGGCGGCTCGCTCACTGAGCGGCGGCGACGAGGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 121 GCCAACTCCATCACTAGGGTTTCCT 145

RESULT 5
ACC58491/C
ID ACC58491 standard; DNA; 145 BP.

XX ACC58491;
XX
XX 26-AUG-2003 (first entry)
XX
XX Adeno associated virus inverted terminal repeat region.
XX AAV; inverted terminal repeat; insect; gene therapy; vector; ss.
XX Adeno associated virus.
XX
XX Key Location/Qualifiers
XX stem_loop 21..145
XX /*tag= a
XX
XX WO2003042361-A2.
XX
XX 22-MAY-2003.
XX
XX 08-NOV-2002; 2002WO-US035829.
XX
XX 09-NOV-2001; 2001US-00986618.
XX 13-AUG-2002; 2002US-00216870.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Kotin RM, Urabe M, Ding C;
XX WPI; 2003-449568/42.
XX
XX Producing an adeno-associated virus (AAV) in an insect cell, e.g. for
XX gene therapy, comprises introducing an insect cell-compatible vector into
XX an insect cell and maintaining the insect cell under conditions that
XX produce AAV.
XX
XX Disclosure; Fig 5A; 84pp; English.

XX The present sequence is that of an adeno associated virus (AAV)
XX palindromic inverted terminal repeat (ITR) region. An AAV ITR is used in
XX novel insect cell-compatible vectors of the invention that are designed
XX for the production of AAV in insect cells. The vectors comprise an AAV

CC ITR and coding sequences for AAV VP1, VP2 and VP3 capsid proteins, for
CC AAV non-structural replication (Rep) proteins Rep52 or Rep40, and for AAV
CC Rep78 or Rep68. The method of producing AAV in insect cells provides an
CC efficient, safe and economical means of producing a large amount of
CC recombinant AAV particles which may be used in gene therapy
XX
SQ Sequence 145 BP; 23 A; 49 C; 52 G; 21 T; 0 U; 0 Other;

Query Match 100.0%; Score 145; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 8.1e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTAGAGCGCGCGACCAAGGTCGCC 60
Db 145 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTAGAGCGCGCGACCAAGGTCGCC 86
QY 61 CGACGCCCGGGCTTTCCCGCGGGCTCAGTGCAGCGGAGCGCGAGAGGGAGTG 120
Db 85 CGACGCCCGGGCTTTCCCGCGGGCTCAGTGCAGCGGAGCGCGAGAGGGAGTG 26
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 25 GCCAACTCCATCACTAGGGGTTCT 1

RESULT 6
ACF35876
ID ACF35876 standard; DNA; 145 BP.

AC ACF35876;
XX
DT 06-NOV-2003 (first entry)
XX
DE AAV-2 left inverted terminal repeat sequence.
XX
KW Adenovirus; Rep78; Rep68; CAP; gene therapy; AAV-2; ITR; ds.
XX
OS Adeno associated virus.
XX
PN WO2003061582-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US0001624.
XX
PR 18-JAN-2002; 2002US-0349532P.
XX
PA (UYDU-) UNIV DUKE.
XX
PI Li C, Zhang X;
XX
PI WPI; 2003-627412/59.
XX
DR
XX
PT New recombinant adenovirus comprising an adenovirus that encodes one or
FT more AAV REP78/68 polypeptides that are inducibly expressed, useful as a
PT vector for gene therapy.
XX
XX Example; Page 120; 122pp; English.

XX The invention relates to a recombinant adenovirus comprising an
CC adenovirus that encodes one or more AAV REP78/68 polypeptides that are
CC inducibly expressed. A complete virus-mediated system for recombinant AAV
CC production is provided which comprises: (a) a first recombinant
CC adenovirus encoding one or more AAV REP78/68 polypeptides and one or more
CC viral capsid polypeptides; (b) a second recombinant adenovirus comprising
CC a gene of interest and AAV inverted terminal repeats that flank the gene
CC of interest; (c) viral helper functions; and (d) a host cell comprising
CC the first recombinant adenovirus, the second recombinant adenovirus, and
CC the viral helper functions. The recombinant AAV is useful as a vector for
CC gene therapy. The present sequence represents an AAV-2 left inverted
CC terminal repeat (ITR) sequence
XX

SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;

Query Match 100.0%; Score 145; DB 8; Length 145;
Best Local Similarity 100.0%; Pred. No. 8.1e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTAGAGCGCGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTAGAGCGCGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTCCCGCGGGCTCAGTGCAGCGGAGCGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTCCCGCGGGCTCAGTGCAGCGGAGCGCGAGAGGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 7
ABX93568
ID ABX93568 standard; DNA; 146 BP.

XX AC ABX93568;
XX
DT 23-MAY-2003 (first entry)
XX
DE Adeno-associated virus-2 inverted terminal repeat.
XX
KW Inverted terminal repeat; Rep recognition sequence; AAV-2; ss; RRS;
XX viral regulatory element.
XX
OS Adeno-associated virus serotype 2.
XX

Key Location/Qualifiers
FT misc_binding 1..41
FT /*tag= a
FT /bound_moiety= "Nucleotides 125-85 of the present
FT sequence"
FT misc_binding 42..50
FT /*tag= b
FT /bound_moiety= "Nucleotides 62-54 of the present
FT sequence"
FT misc_binding 54..62
FT /*tag= c
FT /bound_moiety= "Nucleotides 50-42 of the present
FT sequence"
FT misc_binding 64..72
FT /*tag= d
FT /bound_moiety= "Nucleotides 84-76 of the present
FT sequence"
FT misc_binding 76..84
FT /*tag= e
FT /bound_moiety= "Nucleotides 72-64 of the present
FT sequence"
FT misc_binding 85..125
FT /*tag= f
FT /bound_moiety= "Nucleotides 41-1 of the present sequence"
FT misc_signal 93..108
FT /*tag= g
FT /label= Rep recognition sequence
FT /note= "this RRS is specifically claim in claim 13"
FT misc_recomb 122..127
FT /*tag= h
FT /label= Terminal_resolution_site

US2002182595-A1.

05-DEC-2002.

29-APR-2002; 2002US-00135984.

27-APR-2001; 2001US-0286951P.

XX

FT	repeat_unit	1..20	/tag= b
FT		/rpt_type= INVERTED	
FT		/label= D_repeat	
FT	misc_structure	21..145	
FT		*tag= 1	
FT		/label= T-shaped	
FT		/note= "The naturally occurring 145 bp ITR sequence can form a T-shaped structure for DNA replication when single-stranded"	
FT	repeat_unit	21..62	
FT		*tag= c	
FT		/rpt_type= INVERTED	
FT		/label= A' repeat	
FT		/note= "Forms the stem of a T-shaped structure when base paired with repeat A"	
FT	stem_loop	63..82	
FT		*tag= j	
FT		/note= "Part of T-shaped structure, formed by base pairing between the C' and C repeats"	
FT	repeat_unit	63..70	
FT		*tag= d	
FT		/rpt_type= INVERTED	
FT		/label= C' repeat	
FT	repeat_unit	74..81	
FT		*tag= e	
FT		/rpt_type= INVERTED	
FT		/label= C-repeat	
FT	stem_loop	84..104	
FT		*tag= k	
FT		/note= "Part of T-shaped structure, formed by base pairing between the B' and B repeats"	
FT	repeat_unit	85..92	
FT		*tag= f	
FT		/rpt_type= INVERTED	
FT		/label= B-repeat	
FT	repeat_unit	96..103	
FT		*tag= g	
FT		/rpt_type= INVERTED	
FT		/label= B' repeat	
FT	repeat_unit	104..145	
FT		*tag= h	
FT		/rpt_type= INVERTED	
FT		/label= A-repeat	
FT		/note= "Forms the stem of a T-shaped structure when base paired with A' repeat"	
FT	repeat_unit	146..165	
FT		*tag= i	
FT		/label= D'	
FT		/note= "Additional D' sequence"	
XX			
PN	WO9636364-A1.		
XX			
XX	21-NOV-1996.		
XX			
PP	14-MAY-1996;	96WO-US006786.	
XX			
PR	15-MAY-1995;	95US-00440738.	
XX			
PA	(SAMU/) SAMULSKI R J.		
PA	(XIAO/) XIAO X.		
XX			
PI	Samulski RJ, Xiao X;		
XX			
XX	WPI; 1997-042643/04.		
XX			
PT	Double-D sequence directs adeno-associated virus integration into host genome - used in gene therapy, maintains full length coding sequence of therapeutic gene.		
PT			
XX			
XX	Claim 1; Fig 9; 55pp; English.		
XX			
CC	The adeno associated virus (AAV) has a 145 bp inverted terminal repeat		

(ITR) located at each end of its genome. In addition to being able to base pair with each other, the ITRs can also individually fold back on themselves through the base pairing of A, A', B, B' and C, C' sequences to form a T-shaped structure for DNA replication (see features table). It has been found that viral mutants with deleted D sequences are unable to replicate their DNA. Addition of a second D repeat (D') to the naturally occurring ITR resulted in a sequence, designated double-D, which was sufficient to carry out the functions normally required of two wild-type ITRs during a lytic AAV viral infection, i.e. it is capable of directing replication and assembly into AAV, and/or the integration into the host genome, of recombinant DNA containing the double-D sequence are useful in gene and viral particles containing the double-D sequence are useful in gene therapy. Replication and integration into the host genome is completely effected through the double-D sequences, ensuring that the heterologous gene sequences remain intact

XX
SQ Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;

Query Match 100.0%; Score 145; DB 2; Length 165;
Best Local Similarity 100.0%; Pred.No. 8.2e-30;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGGCACCTCCCTCTCTCTGCGCGCTCGTCTCACTGAGCGCGGCGACCAAGGTCGCG 60
DB 21 TTGGGCACCTCCCTCTCTCTGCGCGCTCGTCTCACTGAGCGCGGCGACCAAGGTCGCG 80
QY 61 CGACGCGCGGGCTTGCCCGCGCGCTCTAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 81 CGACGCGCGGGCTTGCCCGCGCGCTCTAGTGAGCGAGCGCGCGAGAGGGAGTG 140
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 141 GCCAACTCCATCACTAGGGGTTCT 165

RESULT 10
AAD03535
ID AAD03535 standard; DNA; 165 BP.
XX
AC AAD03535;
XX
DT 19-JUN-2001 (first entry)
XX
DE Inverted terminal repeat double DD DNA sequence.
XX
KW Recombinant adeno-associated virus; RAAV; circular intermediate; ITR;
KW inverted terminal repeat; haemostatic; antiskinning; neuroprotective;
KW antianaemic; nootropic; blood disorder; sickle cell anaemia;
KW thalassaemia; neurological disorder; haemophilia; Alzheimer's disease;
KW muscle disorder; Parkinson's disease; gene delivery; erythropoietin; epo;
KW CTR; cystic fibrosis transmembrane conductance receptor; gaucher's disease;
KW tyrosine hydroxylase; glucocerebrosidase gene; gaucher's disease;
KW gene therapy; Double DD; ds.
XX
OS Unidentified.
XX
OS
XX
PN WC200125465-A1.
XX
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027863.
XX
PR 07-OCT-1999; 99US-0158209P.
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (ENGE/) ENGELHARDT J F.
PA (DONG/) DONGSHENG D.
PA (ZIYI/) ZIYING Y.
XX
XX Engelhardt JF, Dongsheng D, Ziyong Y;
XX
XX WPI; 2001-266321/27.
XX

PT Composition for transferring recombinant DNAs and to express a
 PT polypeptide in a host cell, comprises two recombinant adeno-associated
 PT viruses.
 XX
 PS Disclosure; Page 139; 144pp; English.
 XX
 CC The patent discloses a composition comprising at least two recombinant
 CC adeno-associated viruses (rAAVs). Each rAAV contains a recombinant DNA
 CC molecule comprising three DNA segments linked together. The first and the
 CC third DNA segments comprise 5' and 3' inverted terminal repeats (ITRs)
 CC respectively from the circular intermediate of AAV. The second DNA
 CC segment in each virus is different and does not comprise AAV sequence. It
 CC preferably comprises sequences encoding a therapeutically effective
 CC polypeptide such as the cystic fibrosis transmembrane conductance
 CC receptor gene (CFTR), the erythropoietin (epo) gene, the tyrosine
 CC hydroxylase gene (Parkinson's syndrome) or the glucocere- brosidase gene
 CC (Gaucher's disease). The circular intermediate of AAV imparts increased
 CC episomal stability and persistence of the vector in the host cell.
 CC Compositions comprising rAAV sequences are useful for transferring
 CC recombinant DNAs to a host cell and express a polypeptide in a host cell.
 CC The recombinant vector is useful in medical therapy, which includes
 CC treatment or prophylaxis of blood disorders (e.g. sickle cell anaemia,
 CC thalassaemia, haemophilia), neurological disorders, such as Alzheimer's
 CC disease, Parkinson's disease, muscle disorders involving skeletal
 CC cardiac or smooth muscle. AAV vector is used as a delivery vehicle for
 CC gene therapy. The present sequence is the ITR DD DNA sequence referred to
 CC as "double sequence". This sequence is disclosed in U.S. Patent No.
 CC 5,478,745. This sequence is not used in the present invention
 XX
 SQ Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 4; Length 165;
 Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCTCTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 60
 DB 21 TTGGCCTCTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 80
 QY 61 CGACGCCCGGGTTTGGCGCGGCGCTCGCTCACTAGGCGCGGCGAGCGAGAGGGAGTG 120
 DB 81 CGACGCCCGGGTTTGGCGCGGCGCTCGCTCACTAGGCGCGGCGAGCGAGAGGGAGTG 140
 QY 121 GCCAACTCCATCACTAGGGGTTTCCT 145
 DB 141 GCCAACTCCATCACTAGGGGTTTCCT 165
 RESULT 11
 ABS69886
 ID ABS69886 standard; DNA; 165 BP.
 XX
 AC ABS69886;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE Human adeno-associated virus 2 terminal repeat DD sequence.
 XX
 KW Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
 KW adenosine deaminase deficiency; severe combined immune deficiency; PAH;
 KW beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
 KW low density lipoprotein gene; familial hypercholesterolaemia;
 KW hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
 KW phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
 KW dystrophin gene; muscular dystrophy; cystic fibrosis; immunostimulant;
 KW human cystic fibrosis transmembrane conductance regulator gene;
 KW artianaseamic; antilipemic; nootropic; cyostatic; dermatological;
 KW human adeno-associated virus 2; AAV2; terminal repeat; ds.
 XX
 OS Homo sapiens.
 XX
 XX US2002102731-A1.
 PN
 XX

PD 01-AUG-2002.
 XX
 PP 12-FEB-2001; 2001US-00782378.
 XX
 PR 02-OCT-2000; 2000US-0237747P.
 XX
 PA (UINY) UNIV NEW YORK STATE RES FOUND.
 XX
 PI Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
 XX
 XX WPI; 2002-690619/74.
 XX
 DR Producing vector, by introducing vector having nucleotide sequence,
 PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
 PT associated virus terminal repeat, into cell, and culturing cell.
 XX
 PS Disclosure; Page 8; 191pp; English.
 XX
 CC The present invention relates to a new method of producing a vector. The
 CC method involves introducing recombinant vector having nucleotide sequence
 CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
 CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
 CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
 CC end of NS, into cell expressing adenovirus early gene lacking from vector
 CC and culturing cell to produce another vector. The method is useful for
 CC generating vectors, especially mAd vectors. The method is useful in
 CC transferring nucleotide sequences of interest into a cell, for gene
 CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
 CC The nucleotide sequences are useful for treating diseases associated with
 CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
 CC deficiency with severe combined immune deficiency, beta-chain of
 CC haemoglobin gene associated with beta-thalassaemia and sickle cell
 CC disease, receptor for low density lipoprotein gene associated with
 CC familial hypercholesterolaemia, hypoxanthine-guanine
 CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
 CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
 CC dystrophin gene associated with muscular dystrophy, and human cystic
 CC fibrosis transmembrane conductance regulator gene associated with cystic
 CC fibrosis. The present nucleic acid sequence represents a human adeno-
 CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
 CC methods of the invention
 XX
 SQ Sequence 165 BP; 24 A; 59 C; 53 G; 29 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 6; Length 165;
 Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCTCTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 60
 DB 21 TTGGCCTCTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 80
 QY 61 CGACGCCCGGGTTTGGCGCGGCGCTCGCTCACTAGGCGCGGCGAGCGAGAGGGAGTG 120
 DB 81 CGACGCCCGGGTTTGGCGCGGCGCTCGCTCACTAGGCGCGGCGAGCGAGAGGGAGTG 140
 QY 121 GCCAACTCCATCACTAGGGGTTTCCT 145
 DB 141 GCCAACTCCATCACTAGGGGTTTCCT 165
 RESULT 12
 ABV77279/c
 ID ABV77279 standard; DNA; 165 BP.
 XX
 AC ABV77279;
 XX
 DT 28-MAR-2003 (first entry)
 XX
 DE Nucleotide sequence of pTRT, a circular adeno-associated virus.
 XX
 XX Circular adeno-associated virus; cAAV; replication; hairpin;
 KW gene therapy; pTRT; ss.
 KW

XX Adeno associated virus.
 XX WO200297056-A2.
 XX 05-DEC-2002.
 XX 31-MAY-2002; 2002WO-US017324.
 XX 31-MAY-2001; 2001US-0294797P.
 XX 07-AUG-2001; 2001US-0313007P.
 XX (UYRQ) UNIV ROCKEFELLER.
 XX Kaplitt MG, Moussatov S;
 XX WPI; 2003-103706/09.
 XX Production of defective viral vectors for gene therapy that are
 XX completely free of helper viral vectors and helper viruses.
 XX Example 4; Page 53; 69pp; English.
 XX The specification describes a nucleotide sequence capable of directing
 XX circular adeno-associated virus (cAAV) replication. This nucleotide
 XX sequence comprises a loop sequence (TGCCAA) flanked on the 5' and 3'
 XX sides by complementary sequences, where a hairpin structure is formed
 XX between the complementary sequences. The cAAVs are useful in gene
 XX therapy, e.g. to treat an acute medical condition a nucleic acid encoding
 XX a therapeutic protein is inserted into the cAAV. The present sequence
 XX represents pTRT, a cAAV that contains a wild-type circularization point
 XX (the TRR domain), consisting of a single ITR flanked by two D-sequences
 XX
 XX Sequence 165 BP; 26 A; 56 C; 56 G; 27 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 145; DB 7; Length 165;
 XX Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGCGCGCAAGGTCGCC 60
 Db 145 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGCGCGCAAGGTCGCC 86
 QY 61 CGACGCGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGCGCGCGCGAGAGGGAGTG 120
 Db 85 CGACGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGCGCGCGCGAGAGGGAGTG 26
 QY 121 GCCAACTCCATCACTAGGGGTTCT 145
 Db 25 GCCAACTCCATCACTAGGGGTTCT 1
 RESULT 13
 AAD44621
 ID AAD44621 standard; DNA; 207 BP.
 XX
 XX AAD44621;
 XX
 XX 13-DEC-2002 (first entry)
 XX 165 bp enzyme attachment site (EAS) DNA.
 XX Prokaryotic library; candidate protein; nucleic acid modification; NAM;
 XX enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
 XX enzymology; cosmetic research; toxic; environmental safety assessment;
 XX nutrient biology; enzyme attachment site; EAS; ds.
 XX Unidentified.
 XX
 XX WO200266653-A2.
 XX 29-AUG-2002.
 XX

PF 14-DEC-2001; 2001WO-US049058.
 XX
 PR 14-DEC-2000; 2000US-0256163P.
 XX
 PA (XENC-) KENCOR INC.
 XX
 PI Li M, Liu Y;
 XX
 PS WPI; 2002-667068/71.
 DR
 XX
 XX New library of prokaryotic pET-24a expression vectors, host cells or
 PT nucleic acid/protein conjugates, useful for screening candidate proteins
 PT and their nucleic acids or modification enzymes for pharmacogenetic
 PT analysis.
 XX
 XX Disclosure; Fig 50B; 127pp; English.
 XX
 XX The invention relates to methods and compositions for the construction of
 CC prokaryotic libraries expressing candidate proteins and the use of these
 CC libraries to identify candidate proteins and the nucleic acids encoding
 CC them. The invention provides a library of prokaryotic pET-24a vectors
 CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
 CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
 CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
 CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
 CC by the NAM enzyme. The library is used for identifying candidate proteins
 CC and nucleic acids encoding these proteins, in screening for NAM enzymes
 CC with decreased toxicity for the host cells, or in identifying novel or
 CC improved EASs, which may be used for understanding cellular processes or
 CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
 CC (NAP) conjugates are useful in diagnostic assays and in research
 CC including clinical pharmacology, functional genomics, pharmacogenomics,
 CC agricultural chemicals, environmental safety assessment, chemical sensor,
 CC nutrient biology, cosmetic research or enzymology. These may also be used
 CC in in vitro screening techniques and in assays with target molecules. The
 CC present sequence is an enzyme attachment site (EAS) DNA used in the
 CC invention
 XX
 XX Sequence 207 BP; 36 A; 68 C; 68 G; 35 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 145; DB 6; Length 207;
 XX Best Local Similarity 100.0%; Pred. No. 8.3e-30;
 XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGCGCGCAAGGTCGCC 60
 Db 42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAGTGAGCGCGCGCGCAAGGTCGCC 101
 QY 61 CGACGCGCGGCTTTGCGCGGCGCTCAGTGAGCGCGCGCGCGAGAGGGAGTG 120
 Db 102 CGACGCGCGGCTTTGCGCGGCGCTCAGTGAGCGCGCGCGCGAGAGGGAGTG 161
 QY 121 GCCAACTCCATCACTAGGGGTTCT 145
 Db 162 GCCAACTCCATCACTAGGGGTTCT 186
 RESULT 14
 AAD37254
 ID AAD37254 standard; DNA; 955 BP.
 XX
 XX AAD37254;
 XX
 XX 21-AUG-2002 (first entry)
 XX
 XX Adeno-associated virus (AAV) vector plasmid #1.
 XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX
 XX Homo sapiens.
 OS Unidentified.

OS Chimeric.
 XX WO200183695-A2.
 PN 08-NOV-2001.
 XX 27-APR-2001; 2001WO-US013677.
 XX 28-APR-2000; 2000US-0200777P.
 XX (XIAO/) XIAO X.
 XX Xiao X;
 XX WPI; 2002-049342/06.
 XX New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX Example 1; Page 57; 71pp; English.
 XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid containing human dystrophin minigenes, a muscle creatine
 CC kinase (MCK) promoter and a small polyA signal sequence
 XX Sequence 955 BP; 177 A; 307 C; 296 G; 175 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 145; DB 6; Length 955;
 Best Local Similarity 100.0%; Pred. No. 9.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
 Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
 QY 61 CGAGCGCGCGGCTTTGCCCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 120
 Db 61 CGAGCGCGCGGCTTTGCCCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 120
 QY 121 GCCAACTCCATCACTAGGGGTTCTT 145
 Db 121 GCCAACTCCATCACTAGGGGTTCTT 145
 RESULT 15
 AAD37254/C
 ID AAD37254 standard; DNA; 955 BP.
 AC AAD37254;
 XX
 XX 21-AUG-2002 (first entry)
 XX Adeno-associated virus (AAV) vector plasmid #1.
 XX Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 KW Becker muscular dystrophy; ds.
 XX Homo sapiens.
 OS Unidentified.
 OS Chimeric.
 XX

PN WO200183695-A2.
 XX 08-NOV-2001.
 XX 27-APR-2001; 2001WO-US013677.
 XX 28-APR-2000; 2000US-0200777P.
 XX (XIAO/) XIAO X.
 XX Xiao X;
 XX WPI; 2002-049342/06.
 XX New dystrophin minigene for treating Duchenne or Becker muscular
 PT dystrophy comprises an N-terminal domain or modified N-terminal domain,
 PT rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
 PT gene.
 XX Example 1; Page 57; 71pp; English.
 XX The present invention relates to an isolated nucleotide sequence encoding
 CC a dystrophin minigene. The minigene comprises N-terminal or modified N-
 CC terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
 CC domains and cysteine-rich domains of dystrophin or utrophin genes. The
 CC invention also relates to a recombinant adeno-associated virus (AAV)
 CC comprising dystrophin minigene operably linked to an expression control
 CC element. The dystrophin minigene in operable linkage with an expression
 CC control element, in a recombinant adeno-associated virus or retrovirus is
 CC useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
 CC dystrophy (BMD) in a mammalian subject. The present sequence is AAV
 CC vector plasmid containing human dystrophin minigenes, a muscle creatine
 CC kinase (MCK) promoter and a small polyA signal sequence
 XX Sequence 955 BP; 177 A; 307 C; 296 G; 175 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 145; DB 6; Length 955;
 Best Local Similarity 100.0%; Pred. No. 9.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
 Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 896
 QY 61 CGAGCGCGCGGCTTTGCCCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 120
 Db 895 CGAGCGCGCGGCTTTGCCCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 836
 QY 121 GCCAACTCCATCACTAGGGGTTCTT 145
 Db 835 GCCAACTCCATCACTAGGGGTTCTT 811
 Search completed: April 1, 2004, 15:07:23
 Job time : 249.815 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:05:53 ; Search time 1825.39 Seconds
(without alignments)
3442.959 Million cell updates/sec

Title: US-10-620-039-1
Perfect score: 145
Sequence: 1 TTGGCCACTCCCTCTGCG.....CTCCATCACTAGGGTTCCT 145

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671515995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vt.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	145	6	A46401
2	145	100.0	145	14	AA2LTR1
3	145	100.0	145	14	AA2REPORI
4	145	100.0	165	6	AR034135
5	145	100.0	165	6	II6806
6	145	100.0	165	6	AR223306
7	145	100.0	165	6	AX106702
8	145	100.0	165	6	BD18219
9	145	100.0	207	6	AX703496
10	145	100.0	2116	14	AA2LEFT
11	145	100.0	4675	6	AX135805
12	145	100.0	4675	6	AX286292
13	145	100.0	4675	6	AX753252
14	145	100.0	4675	6	BD094552
15	145	100.0	4675	14	AA2CG
16	145	100.0	4679	6	AX282480
17	145	100.0	4679	14	AF043303
18	145	100.0	4680	6	AR028767
19	145	100.0	4680	6	I62303
20	145	100.0	4681	6	BD242774
21	145	100.0	4683	6	BD242775
22	145	100.0	4683	14	AF028704
23	145	100.0	5932	6	AR235457
24	145	100.0	5932	6	AR235457
25	145	100.0	6142	6	AR235461
26	145	100.0	6142	6	AR235461
27	145	100.0	6565	6	AR235454
28	145	100.0	6565	6	AR235454
29	145	100.0	6714	6	AR235459
30	145	100.0	6714	6	AR235459
31	145	100.0	6924	6	AR235462
32	145	100.0	6924	6	AR235462
33	145	100.0	6924	6	AR235463
34	145	100.0	6924	6	AR235464
35	145	100.0	6924	6	AR235464
36	145	100.0	6924	6	AR235464
37	145	100.0	6981	6	AR235460
38	145	100.0	6981	6	AR235460
39	145	100.0	7054	6	AR235456
40	145	100.0	7054	6	AR235456
41	145	100.0	7405	6	AR235455
42	145	100.0	7405	6	AR235455
43	145	100.0	7492	6	AR235458
44	145	100.0	7492	6	AR235458
45	145	100.0	8698	6	AR222044

ALIGNMENTS

RESULT 1	A46401	A46401	Sequence 4 from Patent WO9523867.	145 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	A46401	A46401	Sequence 4 from Patent WO9523867.				
DEFINITION	A46401	A46401	Sequence 4 from Patent WO9523867.				
ACCESSION	A46401	A46401	Sequence 4 from Patent WO9523867.				
VERSION	A46401.1	GI:2300602					
KEYWORDS	unidentified						
SOURCE	unidentified						
ORGANISM	unclassified						
REFERENCE	1 (bases 1 to 145)						
AUTHORS	Densfle,P., Latta,M., Perricaudet,M. and Vigne,E.						
TITLE	INTEGRATIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF						
JOURNAL	Patent: WO 9523867-A 4 08-SEP-1995;						

Pred. No. is the number of results predicted by chance to have a

COMMENT RHONE POULENC ROBER SA (FR)
Other publication AU 1852695 950918
Other publication FR 2716893 950908.

FEATURES
source
1. 145
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 145; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60

Qy 61 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Db 61 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120

Qy 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 2
AA2LTR1 145 bp DNA linear VRL 27-APR-1993
LOCUS
DEFINITION Adeno-associated virus 2 left terminal sequence.
ACCESSION K01624
VERSION K01624.1 GI:209623
KEYWORDS replication; terminal repeat.
SEGMENT 1 of 2
SOURCE Adeno-associated virus 2H
ORGANISM Adeno-associated virus 2H
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 145)
AUTHORS Lusby,E., Fife,K.H. and Berns,K.I.
TITLE Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA
J. Virol. 34 (2), 402-409 (1980)
MEDLINE 80185149
PUBMED 6245271
2 (bases 1 to 145)
AUTHORS Lefebvre,R.B., Riva,S. and Berns,K.I.
TITLE Conformation takes precedence over sequence in adeno-associated virus DNA replication
Mol. Cell. Biol. 4 (7), 1416-1419 (1984)
MEDLINE 85061247
PUBMED 6504049
COMMENT Original source text: Adeno-associated virus 2H DNA, (clone pSM620 [2]), from KB or HeLa cells.
Both [1] and [2] present the opposite strand from the one presented here. The focus of both papers is the method of replication of the virus. [1] notes that the initial tt is present only 30% of the time; it is shortened to t in 50% of the population and missing altogether in 15% of the population. There is further sequence heterogeneity which can be explained by assuming that the terminal 125 bases, which form an imperfect palindrome, are replaced by their inverted complement during replication. [2] found that deletion of the 9 terminal bases on the right and the 113 terminal bases on the left of AAV 2 genome did not stop DNA replication. Further deletion of an 11-base symmetrical sequence (bases 89 to 99) in the right terminal repetition inhibits DNA replication. Substitution of either an 8-base (cagatctg) or 12-base (cgggtagcgcg) symmetrical sequence unrelated to the original 11-base sequence restores DNA replication. All of this can be explained by assuming that the 125 base palindrome mentioned above form a t-shaped secondary structure which provides a primer for DNA polymerase during replication.

FEATURES
Location/Qualifiers

source 1. 145
/organism="Adeno-associated virus 2H"
/mol_type="genomic DNA"
/db_xref="taxon:10805"

ORIGIN 2 bases upsteam of HaeIII site.

Query Match 100.0%; Score 145; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60

Qy 61 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGAGGAGTG 120
Db 61 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGAGGAGTG 120

Qy 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 3
AA2REPORI 145 bp ss-DNA linear VRL 27-APR-1993
LOCUS
DEFINITION Adeno-associated virus origin of replication (genome 3' terminus).
ACCESSION M10681
VERSION M10681.1 GI:209626
KEYWORDS
SOURCE Adeno-associated virus 2H
ORGANISM Adeno-associated virus 2H
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 145)
AUTHORS Berns,K.I., Hauswirth,W.W., Fife,K.H. and Lusby,E.
TITLE Adeno-associated virus DNA replication
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 43 Pt 2, 781-787 (1979)
MEDLINE 80023388
PUBMED 226321
COMMENT Original source text: Adeno associated virus 2H (AAV2 H) DNA.
FEATURES
source 1. 145
/organism="Adeno-associated virus 2H"
/mol_type="genomic DNA"
/db_xref="taxon:10805"

ORIGIN
Query Match 100.0%; Score 145; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 145 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 86

Qy 61 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGAGGAGTG 120
Db 85 CGACGCCCGGGTTTGC CGCGCGGCTTCAGTGAGCGAGCGGCGGAGGAGTG 26

Qy 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 25 GCCAACTCCATCACTAGGGGTTCT 1

RESULT 4
AR034135 165 bp DNA linear PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 1 from patent US 5869305.
ACCESSION AR034135
VERSION AR034135.1 GI:5949740
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

```

Unclassified.
REFERENCE 1 (bases 1 to 165)
AUTHORS Samulski, R. Jude., and Xiao, X.
TITLE Recombinant viral vector system
JOURNAL Patent: US 5869305-A 1 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 5
LOCUS I16806 165 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5478745.
ACCESSION I16806
VERSION I16806.1 GI:1251714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 165)
AUTHORS Samulski, R. J., and Xiao, X.
TITLE Recombinant viral vector system
JOURNAL Patent: US 5478745-A 1 26-DEC-1995;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 6
LOCUS AR223306 165 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 7 from patent US 6436392.
ACCESSION AR223306
VERSION AR223306.1 GI:23331457
KEYWORDS
SOURCE Unknown.

```

```

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 165)
AUTHORS Engelhardt, J. F., and Duan, D.
TITLE Adeno-associated virus vectors
JOURNAL Patent: US 6436392-A 7 20-AUG-2002;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 7
LOCUS AX106702 165 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0125465.
ACCESSION AX106702
VERSION AX106702.1 GI:13922363
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
Unclassified.
REFERENCE 1
AUTHORS Engelhardt, J. F., Dongsheng, D., and Ziying, Y.
TITLE Adeno-associated viruses and uses thereof
JOURNAL Patent: WO 0125465-A 7 12-APR-2001;
University of Iowa Research Foundation (US) ; Engelhardt, John F. (US) ; Dongsheng, Duan (US) ; Ziying, Yan (US)
FEATURES Location/Qualifiers
source 1..165
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="SEQ ID NO:1 of U.S. Patent No. 5,478,745"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 8
LOCUS BD218219 165 bp DNA linear PAT 17-JUL-2003

```



```
JOURNAL Patent: WO 0132711-A 5 10-MAY-2001;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ARKANSAS (US)
FEATURES
  source
    1. .4675
      /organism="Adeno-associated virus 2"
      /mol_type="unassigned DNA"
      /db_xref="taxon:10804"
      321. .2186
      /note="unnamed protein product"
      /codon_start=1
      /protein_id="CAC39573.1"
      /db_xref="GI:14272041"
      /db_xref="RENTREMBL:CAC39573"
      /translation="MPGPELVIVKPSDLDLHPLGIDSFVNVVAKEWELPDPDMD
      MNLEAPLVAELQLQDFLEWRVSKAEPALFFVQFEGESYFNMHVLVETGVKS
      MVLGRFLSQREKLIQIRYIGIEPTLPNFAVTKTRNGAGGKNKVDCEYIPNLLPK
      TQPELQWNTMEQLSACLTERKELVAQHLTHVSQTQEQNKQNPNSDAPVRS
      KTSARYMELVGLWKGITGTEKQWQIQDQASYISFNAASNSRQIKALDNAGKMSL
      TKTAPDVLVGQPPVEDISNRIYKILNGVDPQYAAASVFLGWATKFGKRNITWLFQ
      PATGKTNIAEIAHTVFPYGCNVNTNENPFNDVDKWIWVEEGKMTAKVYESAKA
      ILGSKVRVDKCKSSAQIDPTPIVITSNTMCAVIDNSTTPEHOQPLDRMEKPEL
      TRRDHFGKVTGKEVDFFRNAXKHVVEHEFYVKGAKRPARPSADADISEPKRV
      RESVAQSTDAEASINADRYQNKSRHVMNMLPFCQCRMNQNSNICFTHQOK
      DCLCEFPVSESPVSVVKAYQKLCYIIHIMGKVPDCTACTADLVNVDLDDCIPEQ"
ORIGIN
  Query Match 100.0%; Score 145; DB 6; Length 4675;
  Best Local Similarity 100.0%; Pred. No. 9.3e-24;
  Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 12
AX286292
LOCUS 4675 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0180840.
ACCESSION AX286292
VERSION AX286292.1 GI:17048540
KEYWORDS Adeno-associated virus 2
SOURCE Adeno-associated virus 2
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Raj, K. and Beard, P.M.
TITLES Cytotoxic agents
JOURNAL Patent: WO 0180840-A 1 01-NOV-2001;
BTG INTERNATIONAL LIMITED (GB)
FEATURES
  source
    1. .4675
      /organism="Adeno-associated virus 2"
      /mol_type="unassigned DNA"
      /db_xref="taxon:10804"
ORIGIN
  Query Match 100.0%; Score 145; DB 6; Length 4675;
  Best Local Similarity 100.0%; Pred. No. 9.3e-24;
  Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 14
AX286292
LOCUS 4675 bp DNA linear PAT 27-AUG-2002
DEFINITION Method of transferring gene.
ACCESSION BD094552
VERSION BD094552.1 GI:22640140
KEYWORDS WO 0132899-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4675)
AUTHORS Ueno, T., Matsumura, H., Tanaka, K., Iwasaki, T., Ueno, M., Fujinaga, K.,
Asada, K. and Kato, I.
TITLES Method of transferring gene
JOURNAL Patent: WO 0132899-A 1 10-MAY-2001;
TAKARA SHUZO CO LTD, TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA,
TOMOKO IWASAKI, MIITSUHIRO UENO, KEI FUJINAGA, KIYOZO ASADA, IKUNOSHIN
KATO
COMMENT OS Adeno-associated virus
PN WO 0132899-A/1
PD 10-MAY-2001
PF 23-OCT-2000 WO 2000JP007373
PR 29-OCT-1999 JP 99P 308839
PI TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA, TOMOKO IWASAKI, PI
```

```
QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 13
AX753252
LOCUS 4675 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 7 from Patent EP1310571.
ACCESSION AX753252
VERSION AX753252.1 GI:32166109
KEYWORDS Adeno-associated virus 2
SOURCE Adeno-associated virus 2
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Gao, G., Wilson, J.M. and Alvira, M.
TITLES A method of detecting and/or identifying adeno-associated virus
(JAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 7 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
  source
    1. .4675
      /organism="Adeno-associated virus 2"
      /mol_type="unassigned DNA"
      /db_xref="taxon:10804"
ORIGIN
  Query Match 100.0%; Score 145; DB 6; Length 4675;
  Best Local Similarity 100.0%; Pred. No. 9.3e-24;
  Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 14
BD094552
LOCUS 4675 bp DNA linear PAT 27-AUG-2002
DEFINITION Method of transferring gene.
ACCESSION BD094552
VERSION BD094552.1 GI:22640140
KEYWORDS WO 0132899-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4675)
AUTHORS Ueno, T., Matsumura, H., Tanaka, K., Iwasaki, T., Ueno, M., Fujinaga, K.,
Asada, K. and Kato, I.
TITLES Method of transferring gene
JOURNAL Patent: WO 0132899-A 1 10-MAY-2001;
TAKARA SHUZO CO LTD, TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA,
TOMOKO IWASAKI, MIITSUHIRO UENO, KEI FUJINAGA, KIYOZO ASADA, IKUNOSHIN
KATO
COMMENT OS Adeno-associated virus
PN WO 0132899-A/1
PD 10-MAY-2001
PF 23-OCT-2000 WO 2000JP007373
PR 29-OCT-1999 JP 99P 308839
PI TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA, TOMOKO IWASAKI, PI
```

```

MITSUHIRO UENO.
PI KEI FUJINAGA, KIYOZO ASADA, IKUNOSHIN KATO
PC C12N15/661.A61K35/12.A61K35/76.A61K48/00.C12N5/10.C12N7/01 CC
Method of transferring gene
FH Key Location/Qualifiers
FT source 1..4675 /organism='Adeno-associated virus'
FT Location/Qualifiers
source 1..4675
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 4675;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCATCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
DB 1 TTGGCCATCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
QY 61 CGACGCCCGGGCTTTGGCGCGCGGCGGCTCAGTGACGAGCGGCGGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTGGCGCGCGGCGGCTCAGTGACGAGCGGCGGAGAGGGAGTG 120
QY 121 GCCAATCCATCACTAGGGTTCTCT 145
DB 121 GCCAATCCATCACTAGGGTTCTCT 145

RESULT 15
AA2CG 4675 bp ss-DNA linear VRL 27-APR-1993
LOCUS Adeno-associated virus 2, complete genome.
DEFINITION J01901 M12405 M12468 M12469
ACCESSION J01901.1 GI:209616
VERSION alternative splicing; complete genome; major coat protein.
KEYWORDS Adeno-associated virus 2
SOURCE Adeno-associated virus 2
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 4532 to 4675)
AUTHORS Samulski, R.J., Srivastava, A., Berns, K.I. and Muzyczka, N.
TITLE Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV
JOURNAL Cell 33 (1), 135-143 (1983)
MEDLINE 84282662
PUBMED 6088052
REFERENCE 2 (bases 1 to 4675)
AUTHORS Srivastava, A., Luby, E.W. and Berns, K.I.
TITLE Nucleotide sequence and organization of the adeno-associated virus 2 genome
JOURNAL J. Virol. 45 (2), 555-564 (1983)
MEDLINE 83164299
PUBMED 6300419
COMMENT Original source text: Adeno-associated virus 2 DNA from human HeLa cells.
FEATURES
source 1..4675
/organism='Adeno-associated virus 2'
/mol_type='genomic DNA'
/db_xref='taxon:10804'
repeat_region 1..145
/notes='5' inverted terminal repeat'
misc_feature 42..83
/notes='flip oriented DNA'
mRNA 287..4447
/notes='major coat protein A mRNA (alt.)'
join(321..1906,2228..2252)
/notes='major coat protein A' (alt.)
CDS /codon_start=1
/protein_id='AAA42372.1'
/db_xref='GI:209617'

/translation='MPGFYEIVIKVPSDLGLPGISDSFVNWVAEKEMWELPPDSMD
LNLQAPLTVAEKLQRLTEMRVSKAPEALFVQPEKGSYFHMHLVETTVGKVS
MYLSEQLQIREKLQIRYRGTEPTLNNFAVTKTRNGAGGKGVDCYIPNYLLPK
TQPELQWNTNMEYLSACLNLTERKLVAQHLTHVSTQQRQNKQENQNSDAVIRS
KTSARYMELVGLVDKGTISEKQIQEDQASIIISFNAASRSRQIKALDNAGKIMSL
TKTADYLVGQPPVEDISSNRIYKLELNGYDQYAAASFVFGNATKFGKNTIWLFG
PATTKTINAEIAHTVPGYCVNNTNENPFNDCKVIMWEEGKMTAVIDSGTFF
ILGSKVRVDQKCSAQIDPTPIVTSNTNCAVIDGNSTTFHQQLQDRMFKEFL
TRRLDHDGKVTQBEVKDFFRWAKDHVVEHEFYVKGKAKRPAPSDADISEPRV
RESVAQPSSTDAEASINADRYONKCSRHVGNMLLFPCCOCERNQNSNTCFTRGQK
DCLECFPVSESQPVSVKKAQYKLCYIHHMGKVPDCTACDLVNVLDLDDCIFEQ"
873..4447
/notes='major coat protein A mRNA (alt.)'
join(993..1906,2228..2252)
/notes='major coat protein Aa (alt.)'
/codon_start=1
/protein_id='AAA42373.1'
/db_xref='GI:209618'
/translation='MELVGLVLDKGTSEKOWIQEDQASYSFNAASRSRQIKALD
NAGKIMSLTKTADYLVGQPPVEDISSNRIYKLELNGYDQYAAASFVFGNATKFGK
RTNWLFGPATTKTINAEIAHTVPGYCVNNTNENPFNDCKVIMWEEGKMTA
RNTIWLFGPATTKTINAEIAHTVPGYCVNNTNENPFNDCKVIMWEEGKMTA
KVESAKAILGGSKVRVDQKCSAQIDPTPIVTSNTNCAVIDGNSTTFHQQLQ
DMXFEFLTRLDHDGKVTQBEVKDFFRWAKDHVVEHEFYVKGKAKRPAPSDA
DISEPRVRESVAQPSSTDAEASINADRYONKCSRHVGNMLLFPCCOCERNQNS
ICFTHGQKDCLECFPVSESQPVSVKKAQYKLCYIHHMGKVPDCTACDLVNVLDL
CIFEQ"
993..2186
/notes='major coat protein A' (alt.)'
/codon_start=1
/protein_id='AAA42375.1'
/db_xref='GI:209620'
/translation='MELVGLVLDKGTSEKOWIQEDQASYSFNAASRSRQIKALD
NAGKIMSLTKTADYLVGQPPVEDISSNRIYKLELNGYDQYAAASFVFGNATKFGK
RNTIWLFGPATTKTINAEIAHTVPGYCVNNTNENPFNDCKVIMWEEGKMTA
KVESAKAILGGSKVRVDQKCSAQIDPTPIVTSNTNCAVIDGNSTTFHQQLQ
DMXFEFLTRLDHDGKVTQBEVKDFFRWAKDHVVEHEFYVKGKAKRPAPSDA
DISEPRVRESVAQPSSTDAEASINADRYONKCSRHVGNMLLFPCCOCERNQNS
ICFTHGQKDCLECFPVSESQPVSVKKAQYKLCYIHHMGKVPDCTACDLVNVLDL
CIFEQ"
1853..4447
/notes='major coat protein B mRNA (alt.)'
1907..2227
/notes='major coat protein A intron'
2310..4324
/notes='major coat protein B'
/codon_start=1
/protein_id='AAA42376.1'
/db_xref='GI:209621'
/translation='WATGSGAPMADNNEGADGVNNGSNHWCDSITWMDRVTITSTRT
WALPTYNHLYKQISSQGSANDHYEGYSTPWGYEDNRFCHFSRPMQBLNNNW
GFRKRLNFKLNIQVKEVQNDGTTTIANNTSTVQVETDSEYQLPYLGSAGQGL
PPFADVFMVPGQYLLTNNGSQVRSRFSFLEYTPSQMLKATGNFTPTFYTFDVPF
HSSYASHQSLDRMLNPLIDQYLYLSRNTPTSGTTTQSKLQSQAGASDIRQSRNWL
PGPCYQRQSKTANNNSYSWTGATKHNGLRSDLVNPMASHKDDDEEKFFPSQ
VLIFGQSGKTNVNIETWIDEIEITNTEVATEQYGSVSTNLRQNRQATADVN
TQGLPLPMVMVQDRDVLQGPVWAKIPTHIDGHPSPLMGGFGLGHPPLIILKNTVP
NGLPTTFAAKFASRITQYSTGHGQRGDRVGAEGKQQTLESRSNVHFQLOQVC"
4531..4675
/notes='3' inverted terminal repeat"
4592..4634
repeat_region
misc_feature

```

```

ORIGIN      5' end of genomic DNA.
Query Match      100.0%; Score 145; DB 14; Length 4675;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 145, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGGCCGGGGGACCAAGGTCGCC 60
Db      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGGCCGGGGGACCAAGGTCGCC 60

Qy      61 CGAGCCCGGGCTTTGCCCCGGGGCGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120
Db      61 CGAGCCCGGGCTTTGCCCCGGGGCGCTCAGTGAGCGAGCGCGGCGAGAGGGAGTG 120

Qy      121 GCCAACTCCATCACTAGGGTTTCCT 145
Db      121 GCCAACTCCATCACTAGGGTTTCCT 145

```

Search completed: April 1, 2004, 16:04:09
Job time : 1829.39 secs